

AGES OF YOUNG STAR CLUSTERS, MASSIVE BLUE STRAGGLERS, AND THE UPPER MASS LIMIT OF STARS: ANALYZING AGE-DEPENDENT STELLAR MASS FUNCTIONS

F. R. N. SCHNEIDER¹, R. G. IZZARD¹, S. E. DE MINK^{2,3,11}, N. LANGER¹, A. STOLTE¹, A. DE KOTER^{4,5},
 V. V. GVARAMADZE^{6,7}, B. HUßMANN¹, A. LIERMANN^{8,9}, AND H. SANA^{4,10}

¹ Argelander-Institut für Astronomie der Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany; fschneid@astro.uni-bonn.de

² Observatories of the Carnegie Institution for Science, 813 Santa Barbara St, Pasadena, CA 91101, USA

³ Cahill Center for Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

⁴ Astronomical Institute “Anton Pannekoek”, Amsterdam University, Science Park 904, 1098 XH, Amsterdam, The Netherlands

⁵ Instituut voor Sterrenkunde, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium

⁶ Sternberg Astronomical Institute, Lomonosov Moscow State University, Universitetskij Pr. 13, Moscow 119992, Russia

⁷ Isaac Newton Institute of Chile, Moscow Branch, Universitetskij Pr. 13, Moscow 119992, Russia

⁸ Max Planck Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

⁹ Leibniz Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

¹⁰ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

Received 2013 September 15; accepted 2013 November 13; published 2013 December 16

ABSTRACT

Massive stars rapidly change their masses through strong stellar winds and mass transfer in binary systems. The latter aspect is important for populations of massive stars as more than 70% of all O stars are expected to interact with a binary companion during their lifetime. We show that such mass changes leave characteristic signatures in stellar mass functions of young star clusters that can be used to infer their ages and to identify products of binary evolution. We model the observed present-day mass functions of the young Galactic Arches and Quintuplet star clusters using our rapid binary evolution code. We find that the shaping of the mass function by stellar wind mass loss allows us to determine the cluster ages as 3.5 ± 0.7 Myr and 4.8 ± 1.1 Myr, respectively. Exploiting the effects of binary mass exchange on the cluster mass function, we find that the most massive stars in both clusters are rejuvenated products of binary mass transfer, i.e., the massive counterpart of classical blue straggler stars. This resolves the problem of an apparent age spread among the most luminous stars exceeding the expected duration of star formation in these clusters. We perform Monte Carlo simulations to probe stochastic sampling, which support the idea of the most massive stars being rejuvenated binary products. We find that the most massive star is expected to be a binary product after 1.0 ± 0.7 Myr in Arches and after 1.7 ± 1.0 Myr in Quintuplet. Today, the most massive 9 ± 3 stars in Arches and 8 ± 3 in Quintuplet are expected to be such objects. Our findings have strong implications for the stellar upper mass limit and solve the discrepancy between the claimed $150 M_{\odot}$ limit and observations of four stars with initial masses of $165\text{--}320 M_{\odot}$ in R136 and of supernova 2007bi, which is thought to be a pair-instability supernova from an initial $250 M_{\odot}$ star. Using the stellar population of R136, we revise the upper mass limit to values in the range $200\text{--}500 M_{\odot}$.

Key words: binaries: general – blue stragglers – open clusters and associations: individual (Arches, Quintuplet) – stars: luminosity function, mass function – stars: mass-loss

Online-only material: color figures

1. INTRODUCTION

Massive stars play a key role in our universe. They drive the chemical evolution of galaxies by synthesizing most of the heavy elements. Their strong stellar winds, radiation feedback, powerful supernova explosions, and long gamma-ray bursts shape the interstellar medium. They are thought to have played an essential role in reionizing the universe after the dark ages and are visible up to large distances.

Unfortunately, our understanding of the formation and evolution of the most massive stars in the local universe is incomplete (Langer 2012). Recently, it was established that most of the massive stars in the Milky Way are actually part of a binary star system and that more than 70% of them will exchange mass with a companion during their life (Sana et al. 2012). Our understanding of these stars is further hampered by two major controversies. The first one, the cluster age problem, concerns the ages of the youngest star clusters. Emerging star clusters are

expected to form stars in a time span shorter than the lifetime of their most massive members (Elmegreen 2000; Kudryavtseva et al. 2012). In contrast, the most luminous stars in two of the richest young clusters in our Galaxy, the Arches and Quintuplet clusters, show an apparently large age range. Their hydrogen- and nitrogen-rich Wolf–Rayet (WNh) stars appear significantly younger than most of their less luminous O stars (Martins et al. 2008; Liermann et al. 2012). Similar age discrepancies are observed in other young stellar systems (Massey 2003), such as the Cyg OB2 association (Herrero et al. 1999; Gvaramadze & Bomans 2008a; Negueruela et al. 2008) and the star clusters Pismis 24 (Gvaramadze et al. 2011) and NGC 6611 (Hillenbrand et al. 1993; Gvaramadze & Bomans 2008b).

The second controversy, the maximum stellar mass problem, concerns the stellar upper mass limit. Such an upper mass limit is theoretically motivated by the Eddington limit, which may prevent stellar mass growth by accretion above a certain mass (Larson & Starrfield 1971). Observationally, an upper mass limit of about $150 M_{\odot}$ is derived from the individual stellar mass distributions of the Arches and the R136 clusters (Weidner

¹¹ Einstein Fellow.

& Kroupa 2004; Figer 2005; Koen 2006) and from a broader analysis of young stellar clusters (Oey & Clarke 2005). This result is questioned by a recent analysis of very massive stars in the core of R136, in which stars with initial masses of up to about $320 M_{\odot}$ are found (Crowther et al. 2010). Furthermore, recently detected ultraluminous supernovae in the local universe are interpreted as explosions of very massive stars (Langer et al. 2007); for example, SN 2007bi is well explained by a pair-instability supernova from an initially $250 M_{\odot}$ star (Gal-Yam et al. 2009; Langer 2009).

Here we show that both controversies can be resolved by considering a time-dependent stellar mass function in young star clusters that accounts for stellar wind mass loss and binary mass exchange. We perform detailed population synthesis calculations of massive single and binary stars that include all relevant physical processes affecting the stellar masses and compare them to the observed present-day mass functions of the Arches and Quintuplet clusters. Our methods and the observations of the mass functions of the Arches and Quintuplet clusters are described in Section 2. We analyze the Arches and Quintuplet clusters in Section 3 to derive cluster ages and identify possible binary products by fitting our models to the observed mass functions. Stochastic sampling effects are investigated in Section 4, and the implications of our findings for the upper mass limit are explored in Section 5. We discuss our results in Section 6 and give final conclusions in Section 7.

2. METHODS AND OBSERVATIONAL DATA

We analyze the Arches and Quintuplet clusters in two steps. First, we model their observed stellar mass functions to determine, e.g., the initial mass function (IMF) slopes and the cluster ages. We set up a dense grid of single and binary stars, assign each stellar system in the grid a probability of existence given the initial distribution functions (see Section 2.2), and evolve the stars in time using our rapid binary evolution code (see Section 2.1). Present-day mass functions are then constructed from the individual stellar masses at predefined ages. This ensures that all the relevant physics such as stellar wind mass loss and binary mass exchange, which directly affects stellar masses, is factored in our mass functions.

Second, we investigate stochastic sampling effects to compute, e.g., the probability that the most massive stars in the Arches and Quintuplet clusters are binary products. To that end, we randomly draw single and binary stars from initial distribution functions until the initial cluster masses are reached and, again, evolve the drawn stellar systems with our rapid binary evolution code. The setup of these Monte Carlo experiments is described in detail in Section 2.3.

The initial distribution functions used in the above-mentioned steps are summarized in Section 2.2, and an overview of the observations of the Arches and Quintuplet clusters to which we compare our models is given in Section 2.4. We bin mass functions in a nonstandard way to compare them to observations; the binning procedure is described in Section 2.5.

2.1. Rapid Binary Evolution Code

The details of our population synthesis code are described in Schneider et al. (2013) and de Mink et al. (2013a). Here we briefly summarize the most important methods and assumptions that are used to derive our results.

We use a binary population code (Izzard et al. 2004, 2006, 2009) to evolve single and binary stars and follow the evolution

of the stellar masses and of other stellar properties as a function of time. Our code is based on a rapid binary evolution code (Hurley et al. 2002) that uses analytic functions (Hurley et al. 2000) fitted to stellar evolutionary models with convective core overshooting (Pols et al. 1998) to model the evolution of single stars across the whole Hertzsprung–Russell diagram. We use a metallicity of $Z = 0.02$.

Stellar wind mass loss (Nieuwenhuijzen & de Jager 1990) is applied to all stars with luminosities larger than $4000 L_{\odot}$ (Hurley et al. 2000). The mass accretion rate during mass transfer is limited to the thermal rate of the accreting star (Wellstein et al. 2001). Binaries enter a contact phase and merge if the mass ratio of accretor to donor is smaller than a critical value at the onset of Roche lobe overflow (RLOF; de Mink et al. 2013a). When two main-sequence (MS) stars merge, we assume that 10% of the total mass is lost and that 10% of the envelope mass is mixed with the convective core (de Mink et al. 2013a).

Photometric observations of star clusters cannot resolve individual binary components. In order to compare our models to observations we assume that binaries are unresolved in our models and determine masses from the combined luminosity of both binary components utilizing our mass–luminosity relation. Hence, unresolved, preinteraction binaries contribute to our mass functions.

We concentrate on MS stars because stars typically spend about 90% of their lifetime in this evolutionary phase; moreover, our sample stars used for comparison are observationally color selected to remove post-MS objects. If a binary is composed of a post-MS and an MS star, we take only the MS component into account.

2.2. Initial Distribution Functions for Stellar Masses and Orbital Periods

We assume that primary stars in binaries and single stars have masses M_1 distributed according to a power-law IMF with slope Γ ,

$$\xi(M_1) = \frac{dN}{dM_1} = A M_1^{\Gamma-1}, \quad (1)$$

in the mass range $1\text{--}100 M_{\odot}$ (where A is a normalization constant). Secondary star masses M_2 are taken from a flat mass-ratio distribution; that is, all mass ratios $q = M_2/M_1 \leq 1$ are equally probable (Sana et al. 2012). The initial orbital periods for binaries with at least one O star, i.e., a primary star with $M_1 \gtrsim 15 M_{\odot}$, mass ratio $q \geq 0.1$, and a period $0.15 \leq \log(P/\text{days}) \leq 3.5$, are taken from the distribution of stars in Galactic open clusters (Sana et al. 2012). The initial periods of all other binaries follow a flat distribution in the logarithm of the orbital period (Öpik 1924). Orbital periods are chosen such that all interacting binaries are taken into account; that is, the maximum initial orbital separation is $10^4 R_{\odot}$ (≈ 50 AU). Binaries with wider orbits would be effectively single stars.

2.3. Monte Carlo Experiments

To address stochastic sampling, we perform Monte Carlo simulations of star clusters and investigate the probability that the most massive star in a star cluster is a product of binary evolution as a function of cluster mass, age, binary fraction, and IMF slope.

We assume that all stars are coeval and that every star cluster forms from a finite supply of mass, with stellar masses

stochastically sampled from initial distribution functions. While single stars are sampled from the IMF, binary stars are chosen from a larger parameter space of primary and secondary masses and orbital periods. This larger parameter space is better sampled in clusters of higher mass.

We draw single and binary stars for a given binary fraction from the initial distribution functions of primary mass, secondary mass, and orbital period until a given initial cluster mass M_{cl} is reached. Here we consider only stars with masses in the range of $1\text{--}100 M_{\odot}$. The true cluster masses are therefore larger if stars below $1 M_{\odot}$ are added. Including these stars according to a Kroupa IMF (Kroupa 2001) increases the true cluster mass by 20% and 89% for high-mass ($\geq 1 M_{\odot}$) IMF slopes of $\Gamma = -0.70$ and $\Gamma = -1.35$, respectively.

After the stellar content of a cluster is drawn, we evolve the stars in time to analyze whether the most massive stars at a given cluster age result from close binary interaction. Repeating this experiment 1000 times provides the probability that the most massive cluster star formed from binary interactions, how long it takes, on average, until the most massive star is a product of binary evolution (τ_{B}), and how many stars have, on average, a mass larger than that of the most massive cluster star (M_{S}) that did not accrete from a companion, ($N(M > M_{\text{S}})$). The most massive star that did not accrete from a companion can be a genuine single star or a star in a binary where binary mass transfer has not yet happened. From here on we refer to this star as “the most massive genuine single star.” We evolve and distribute stars as described in Sections 2.1 and 2.2.

To compare our Monte Carlo simulations with observations of the Arches and Quintuplet clusters, we need to know the corresponding cluster masses M_{cl} in our Monte Carlo experiments. We use IMF slopes of $\Gamma = -0.7$ as later determined in Section 3.1 for both clusters. The observations used for comparison (Stolte et al. 2005; Hußmann et al. 2012) are complete for masses $> 10 M_{\odot}$, corresponding to 234 and 134 stellar systems in the Arches and Quintuplet and integrated masses of stars more massive than $10 M_{\odot}$ of 7200 and $3100 M_{\odot}$, respectively.

Our best-fitting Monte Carlo models of the central regions of the Arches and Quintuplet with primordial binary fractions of 100% and 60% have 225 ± 10 stellar systems with an integrated (initial) mass of $(7993 \pm 361) M_{\odot}$ and 136 ± 10 stellar systems with an integrated initial mass of $(3240 \pm 244) M_{\odot}$, respectively. These models correspond to initial cluster masses of $M_{\text{cl}} \approx 1.5 \times 10^4 M_{\odot}$ and $M_{\text{cl}} \approx 0.9 \times 10^4 M_{\odot}$, respectively, in stars with $1 \leq M/M_{\odot} \leq 100$.

We assume that binaries are resolved in our Monte Carlo calculations, contrary to when we model mass functions in order to compare to observed mass functions. This is because we make theoretical predictions and are thus interested in individual masses of all stars regardless of whether or not they are in a binary.

2.4. Observations

The observed present-day mass functions of the Arches and Quintuplet clusters were obtained from NAOS/CONICA (NACO) photometry at the Very Large Telescope. The Arches cluster was observed in 2002 over a field of view (FOV) of $27''$. The center of the Quintuplet cluster was imaged with NACO over a FOV of $40''$ in 2003 and 2008, which allowed the construction of a membership source list from proper motions. Both data sets were obtained in the H ($\lambda_c = 1.66 \mu\text{m}$) and K_s ($\lambda_c = 2.18 \mu\text{m}$) passbands. The color information is used to remove likely blue foreground interlopers, red clump, and

giant stars towards the Galactic center line of sight. Details can be found in Stolte et al. (2005) for the Arches cluster and in Hußmann et al. (2012) for the Quintuplet cluster.

In the case of the Arches cluster, the known radial variation of the extinction is removed prior to individual mass determination (Stolte et al. 2002) by employing the extinction law of Rieke & Lebofsky (1985). Masses are then derived from the K_s magnitudes of each star by comparison with a 2 Myr Geneva isochrone (Lejeune & Schaerer 2001). In the case of Quintuplet, the better photometric performance allowed all sources to be individually dereddened to a 4 Myr Padova MS isochrone (Marigo et al. 2008, and references therein) using the recently updated near-infrared extinction law toward the Galactic center line of sight (Nishiyama et al. 2009). As detailed in Hußmann et al. (2012), isochrone ages of 3 and 5 Myr do not significantly alter the shape and slope of the constructed mass function. All mass determinations are based on solar metallicity evolution models.

With the aim to minimize any residual field contamination, only the central $r < 10''$, or 0.4 pc , of the Arches and $r < 12''.5$, or 0.5 pc , of Quintuplet (at an assumed distance of 8.0 kpc to the Galactic center; Ghez et al. 2008) were selected to construct the mass functions. For the Arches cluster, this radial selection corresponds approximately to the half-mass radius, which implies that the mass projected into this annulus is of the order of $\sim 10^4 M_{\odot}$ (see Espinoza et al. 2009; Clarkson et al. 2012; Habibi et al. 2013). In the Quintuplet cluster, the total mass is estimated to be $6000 M_{\odot}$, within the considered 0.5 pc radius (Hußmann et al. 2012). The mass functions in the central regions of Arches and Quintuplet have slopes that are flatter than the usual Salpeter slopes, most likely because of mass segregation (Harfst et al. 2010; Habibi et al. 2013).

The most massive stars in the Arches and Quintuplet clusters are hydrogen- and nitrogen-rich WNh stars. As reliable masses cannot be derived for these WR stars from photometry alone and as several of the WRs in Quintuplet suffered from saturation effects, the most massive stars are excluded from the mass functions. This affects six WNhs stars with uncertain masses in Arches and three (plus seven post-MS, carbon-rich WR stars) in Quintuplet. These WNh stars are expected to contribute to the high-mass tail of the Arches and Quintuplet mass functions.

2.5. Binning Procedure of Mass Functions

Following Stolte et al. (2005), we employ a binning procedure that renders the observed mass functions independent of the starting point of the bins. We shift the starting point by one-tenth of the bin size and create mass functions for each of these starting points. We use a fixed bin size of 0.2 dex to ensure that the number of stars in each bin is not too small and does not introduce a fitting bias (Maíz Apellániz & Úbeda 2005). Each of these 10 mass functions with different starting points is shown when we compare our mass functions to observations.

This procedure results in lowered number counts in the highest-mass bins because only the most massive stars will fall into these bins, as seen in the power-law mass function (black dotted lines in Figure 2) where a kink is visible around $\log M/M_{\odot} \approx 1.85$ (left panel) and $\log M/M_{\odot} \approx 1.45$ (right panel; compare to the convolution of a truncated horizontal line with a box function with the width of the bins). This kink is caused by the binning procedure. Importantly, the observations, our models, and the power-law mass functions in Figures 2 and 8 are binned identically to render the mass functions comparable.

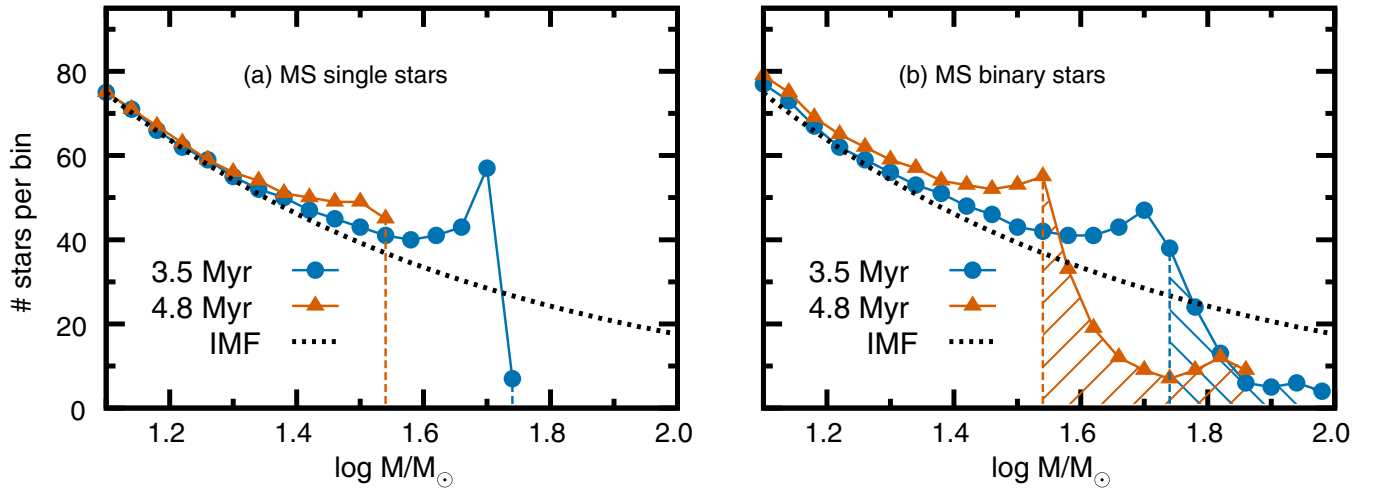


Figure 1. Stellar mass functions, i.e., number of stars per logarithmic stellar mass bin, predicted by our population synthesis models for MS (a) single and (b) binary stars. Circles and triangles show the mass function at 3.5 and 4.8 Myr, respectively. The black dotted line shows the adopted initial mass function ($\Gamma = -0.7$). The peaks in the mass functions caused by stellar wind mass loss are apparent in both plots at about $32 M_{\odot}$ ($\log M/M_{\odot} \approx 1.5$) and $50 M_{\odot}$ ($\log M/M_{\odot} \approx 1.7$), respectively. The tail of stars affected by binary evolution in (b) is highlighted by the hatched regions. The tail extends to about twice the maximum mass expected from single-star evolution, which is indicated by the vertical dashed lines.

(A color version of this figure is available in the online journal.)

3. ANALYSES OF THE ARCHES AND QUINTUPLET CLUSTERS

For a meaningful comparison of the modeled to the observed mass functions, the star cluster and the observations thereof need to fulfill certain criteria. They should be (1) between ~ 2 Myr and ~ 10 Myr in age such that the wind mass loss peak in the mass function is present (see below and Schneider et al. 2013), (2) massive enough such that the mass function samples the largest masses, and (3) homogeneously analyzed, with a complete present-day mass function above $\sim 10 M_{\odot}$. Both the Arches and Quintuplet clusters fulfill all criteria and are therefore chosen for our analysis.

Other possible star clusters, which can be analyzed in principle, are the Galactic center cluster, NGC 3603 YC, Westerlund 1, and R136 in the Large Magellanic Cloud. Trumpler 14 and Trumpler 16 in the Galactic Carina nebula are not massive enough and rather are an OB star association with stars of different ages. For Westerlund 1 (Lim et al. 2013) and NGC 3603 YC (Pang et al. 2013), present-day mass functions were recently derived. A brief inspection of these results shows that both clusters may be suitable for an analysis as performed here for the Arches and Quintuplet clusters. We will investigate this further in the near future. Possibly, NGC 3603 YC is too young, such that its mass function is not yet altered enough by stellar evolution to apply our analysis.

3.1. The Arches and Quintuplet Mass Functions

The initially most massive stars in a cluster end their life first. This depopulates the high-mass end of the stellar mass function. Before that, however, massive stars lose a significant fraction of their initial mass because of strong stellar winds; for example, our $100 M_{\odot}$ star at solar metallicity loses about $40 M_{\odot}$ during core hydrogen burning. Stellar wind mass loss shifts the top of the mass function toward lower masses, and a peak accumulates near its high-mass end (Figures 1(a) and (b)). The location of the peak depends strongly on the cluster age and provides a clock to age-date a star cluster.

Stars in close binary systems exchange mass with their companion either by mass transfer or in a stellar merger. A

fraction of stars gain mass, producing a tail at the high-mass end of the mass function (hatched regions in Figure 1(b)) that extends beyond the most massive single stars (Figure 1(a)) by up to a factor of about 2. The mass gainers appear younger than genuine single stars because their convectively mixed stellar core grows upon mass accretion and mixes fresh fuel into their center, thereby turning their clock backward (van Bever & Vanbeveren 1998). Furthermore, the most massive gainers reach masses that, when interpreted as single stars, have lifetimes that are shorter than the cluster age—they are the massive counterpart of classical blue straggler stars (Schneider et al. 2013).

The mass functions of the cores of the Arches ($r \lesssim 0.4$ pc) and Quintuplet ($r \lesssim 0.5$ pc) clusters (Stolte et al. 2005; Hußmann et al. 2012) reveal both the stellar wind mass loss peak and the tail because of binary mass exchange. Compared to a power law, we find that the Arches and Quintuplet mass functions are overpopulated in the ranges 32 – $50 M_{\odot}$ ($\log M/M_{\odot} = 1.5$ – 1.7) and 20 – $32 M_{\odot}$ ($\log M/M_{\odot} = 1.3$ – 1.5), respectively (Figure 2).

These peaks are well reproduced by our models (Figures 1 and 2). We can thus determine the cluster age because among the stars in the peak are the initially most massive stars that are (1) still on but about to leave the MS and (2) unaffected by binary interactions (we refer to them as turn-off stars in analogy to their position close to the turn-off in a Hertzsprung–Russell diagram). The majority of stars in the peak are turn-off stars, but there are small contributions from unresolved and postinteraction binaries (see Schneider et al. 2013). A correction for wind mass loss then reveals the initial mass of the turn-off stars and hence the age of the cluster.

We can correct for wind mass loss by redistributing the number of excess stars in the peak N such that the mass function is homogeneously filled for masses larger than those of the peak stars up to a maximum mass, the initial mass of the turn-off stars $M_{\text{to},i}$. The number of excess stars is then

$$N = \int_{M_{\text{to},p}}^{M_{\text{to},i}} \xi(M) dM, \quad (2)$$

where $M_{\text{to},p}$ is the present-day mass of the turn-off stars, which can be directly read off from the higher-mass end of the peak,

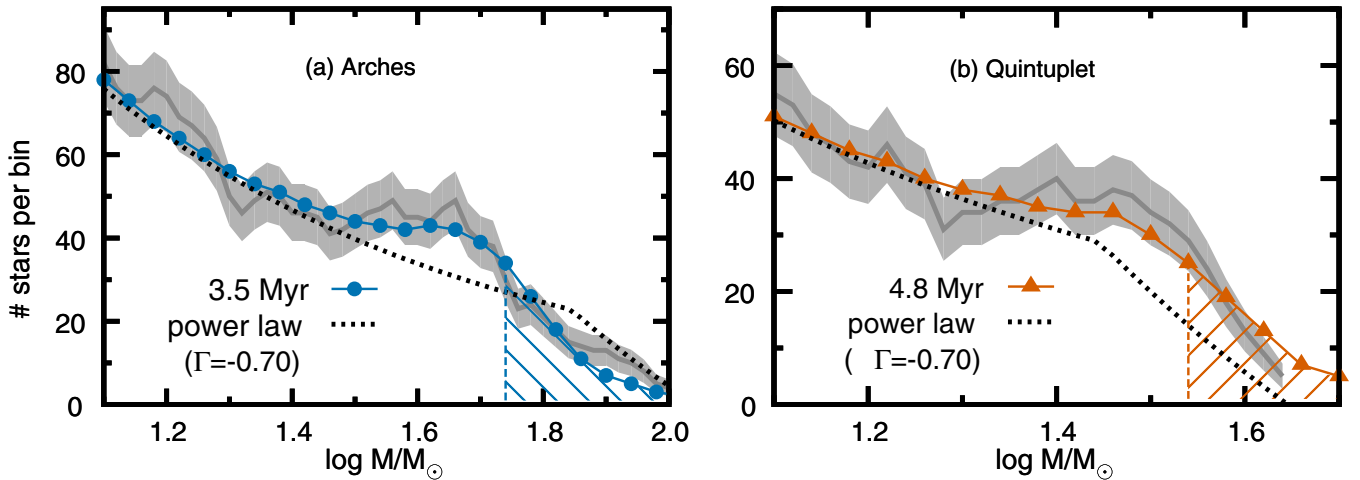


Figure 2. Observed stellar mass functions for a bin size of 0.2 dex of (a) the Arches cluster (Stolte et al. 2005) compared to our 3.5 Myr binary population model from Figure 1(b) (primordial binary fraction 100%) and (b) the Quintuplet cluster (Hußmann et al. 2012) compared to our 4.8 Myr model (primordial binary fraction 60%). The peak and tail of the mass functions are well reproduced by our models. Gray shaded regions around the observed mass functions give Poisson uncertainties and show that the observed peaks deviate by 1σ – 2σ from the power-law mass functions. We plot simple power-law mass functions as dotted lines, binned in the same way as the observations and our models (Section 2.5). The kinks in the power-law mass functions result from our binning procedure.

(A color version of this figure is available in the online journal.)

and $\xi(M)$ is the IMF as defined in Equation (1). The initial mass of the turn-off stars $M_{\text{to},i}$ and hence the cluster age follows from integrating Equation (2),

$$M_{\text{to},i} = \left(\frac{N\Gamma}{A} + M_{\text{to},p}^\Gamma \right)^{1/\Gamma}. \quad (3)$$

The normalizations A of the mass functions to be filled up with the excess stars N (dotted, power-law functions in Figure 2) are $A = 964$ and $A = 639$ for Arches and Quintuplet, respectively, with slopes of $\Gamma = -0.7$ in both cases (see discussion below for why the mass functions are so flat). It is difficult to read off the exact value of $M_{\text{to},p}$ from the observed mass functions because of the binning. However, from binning our modeled mass functions in the same way as the observations, we know that $M_{\text{to},p}$ corresponds to the mass shortly after the peak reached its local maximum (see the vertical dashed lines in Figures 1 and 2). Depending on the exact value of $M_{\text{to},p}$, $\log M_{\text{to},p}/M_\odot = 1.70$ – 1.74 in Arches and $\log M_{\text{to},p}/M_\odot = 1.50$ – 1.54 in Quintuplet, we find 12–14 and 7–10 excess stars in the peaks of the Arches and Quintuplet mass functions, respectively. These numbers of excess stars result in turn-off masses $M_{\text{to},i}$ of 62 – $72 M_\odot$ and 36 – $43 M_\odot$ and hence ages of 3.8 – 3.5 and 5.2 – 4.7 Myr for the Arches and Quintuplet clusters, respectively. These are only first, rough age estimates that will be refined below, and their ranges stem from the uncertainty in reading off $M_{\text{to},p}$ from the observed mass functions.

From the difference between the initial and present-day masses of the turn-off stars in Arches and Quintuplet, we can directly measure the amount of mass lost by these stars through stellar winds. The turn-off stars in Arches lost about 12 – $17 M_\odot$, and the turn-off stars in Quintuplet lost about 4 – $8 M_\odot$ during their MS evolution. This is a new method to measure stellar wind mass loss that does not require measurements of stellar wind mass loss rates and can therefore be used to constrain these.

More accurately, we determine the ages of the Arches and Quintuplet clusters by fitting our population synthesis models (Section 2) to the observed mass functions. First, we fit power-

law functions to the observed mass functions in mass regimes in which they are observationally complete and not influenced by stellar wind mass loss ($10 \lesssim M/M_\odot \lesssim 32$ and $10 \lesssim M/M_\odot \lesssim 20$, respectively). Binary effects are also negligible because stars with such masses are essentially unevolved at the present cluster ages. This fit gives the normalization and a first estimate of the slope of the mass function. We then vary the mass function slope, the cluster age, and the primordial binary fraction in our models simultaneously such that the least squares deviation from the observations is minimized. Our best-fit models are shown in Figure 2 together with the observed mass functions.

We find slopes of $\Gamma = -0.7$, ages of 3.5 ± 0.7 and 4.8 ± 1.1 Myr, and primordial binary fractions of 100% and 60% for the Arches and Quintuplet clusters, respectively. The binary fractions are less robust and may be the same within uncertainties because we do not take the uncertain masses of the WNh stars into account. While our mass function fits contribute to the age uncertainties by only ± 0.3 Myr, its major part, ± 0.6 and ± 1.1 Myr for Arches and Quintuplet, respectively, is due to observational uncertainties in stellar masses of $\pm 30\%$ (Section 6).

Massive stars tend to sink toward the cluster cores because of dynamical friction (mass segregation), thereby flattening the mass function of stars in the core. The derived mass function slopes of $\Gamma = -0.7$ are flatter than the typical Salpeter slope of $\Gamma = -1.35$ (Salpeter 1955) because we investigate only the mass-segregated central regions of both clusters (see Habibi et al. 2013, as well as Section 6.3), i.e., a subsample of stars biased toward larger masses.

In our models (Figure 1), the tail of the Arches mass function contains about 30% unresolved, preinteraction binaries with $\log M/M_\odot \geq 1.76$ ($M \approx 58 M_\odot$) and about 20% with $\log M/M_\odot \geq 1.80$ ($M \approx 63 M_\odot$). For Quintuplet, the fraction of unresolved, preinteraction binaries is about 20% with $\log M/M_\odot \geq 1.56$ ($M \approx 36 M_\odot$) and about 10% with $\log M/M_\odot \geq 1.60$ ($M \approx 40 M_\odot$). The binary fraction among the rejuvenated binary products in the tails is about 55% in our Arches model and 70% in our Quintuplet model, where the remaining stars are single-star binary products, i.e., merger stars.

3.2. The Ages of Arches and Quintuplet

Previously estimated ages for the Arches and Quintuplet clusters lie in the range 2–4.5 Myr (Blum et al. 2001; Figer et al. 2002; Martins et al. 2008) and 2–5 Myr (Figer et al. 1999; Liermann et al. 2010, 2012), respectively. Within these ranges, the age discrepancy between the most luminous cluster members, the WN and the less luminous O stars, accounts for about 1 and 1.5 Myr, respectively (Martins et al. 2008; Liermann et al. 2012), which is eliminated by our method. Our ages of 3.5 ± 0.7 and 4.8 ± 1.1 Myr for the Arches and Quintuplet clusters, respectively, agree with the ages derived from the O stars and dismiss the proposed younger ages from the brightest stars as a result of neglecting binary interactions. The most famous member of the Quintuplet cluster, the Pistol Star, is such an example because it appears to be younger than 2.1 Myr assuming single-star evolution (Figer et al. 1998).

4. STOCHASTIC SAMPLING OF BINARY POPULATIONS

The initial mass of the primary star, the mass ratio, and the orbital period of a binary system determine when mass transfer starts, with more massive and/or closer binaries interacting earlier. Stochastic effects caused by the limited stellar mass budget prevent the formation of all possible binaries in a stellar cluster, i.e., binaries with all possible combinations of primary mass, mass ratio, and orbital period. The likelihood that a binary in a given cluster interacts, e.g., after 2 Myr and that the binary product becomes then the most massive star depends thus on the number of binary stars in that cluster and hence on the total cluster mass. Using Monte Carlo simulations, we investigate the influence of stochastic sampling and binary evolution on the most massive stars in young star clusters (see Section 2.3).

The Galactic star cluster NGC 3603 YC contains NGC 3603-A1, a binary star with component masses $(116 \pm 31) M_\odot$ and $(89 \pm 16) M_\odot$ in a 3.77 day orbit (Schnurr et al. 2008). An initially $120+90 M_\odot$ binary in a 3.77 day orbit starts mass transfer ~ 1.4 Myr after its birth according to the nonrotating models of Ekström et al. (2012). This is the time needed for the $120 M_\odot$ primary star to fill its Roche lobe as a result of stellar evolutionary expansion. This time provides an upper age estimate for NGC 3603 YC. After mass transfer, the secondary star will be the most massive star in the cluster. Were NGC 3603-A1 in a closer orbit, it could already be a binary product today.

To find the probability that the most massive star in a cluster of a given age is a binary product, we investigate how many close binaries are massive enough to become the most massive star by mass transfer. Were the cluster a perfect representation of the initial stellar distribution functions, we could use these functions to derive the probability directly. However, the finite cluster mass and hence sampling density must be considered for comparison with real clusters. Returning to the example of NGC 3603 YC, if the cluster had larger total mass, its binary parameter space would be better sampled, and its most massive star might already be a binary interaction product. For perfect sampling, i.e., infinite cluster mass, the time until a binary product is the most massive star tends toward zero.

The idea that the most massive star in a star cluster may be a binary product resulted from the first discovery of blue straggler stars (Sandage 1953). It was proposed that blue stragglers might stem from binary mass transfer and/or stellar

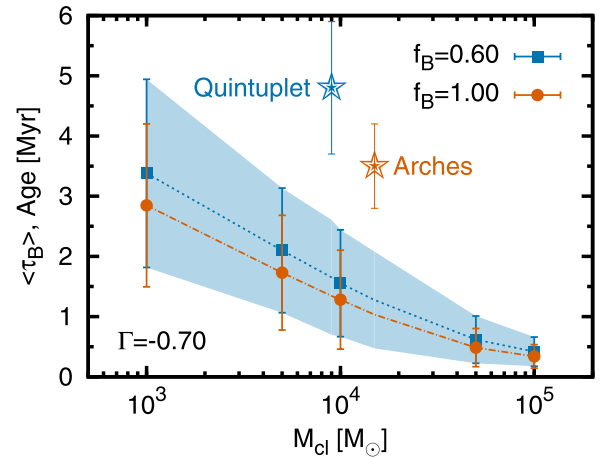


Figure 3. Average time $\langle \tau_B \rangle$ until the most massive star in a star cluster is a product of close binary evolution as a function of cluster mass for two primordial binary fractions f_B and a mass function slope $\Gamma = -0.7$. For steeper, Salpeter-like mass functions, see Figure 4. The error bars are the standard deviation of 1000 realizations of each cluster. The stars indicate the age and central cluster mass of Arches and Quintuplet as derived in this work.

(A color version of this figure is available in the online journal.)

collisions (McCrea 1964; Hills & Day 1976). Stellar population synthesis computations including binary stars then showed that this is indeed possible (e.g., Collier & Jenkins 1984; Pols & Marinus 1994; van Bever & Vanbeveren 1998; Hurley et al. 2001; Chen & Han 2009). Here, we show, using the binary distribution functions of Sana et al. (2012), that the formation of blue stragglers by binary interactions prevails up to the youngest and most massive clusters, and we quantify it for the Arches and Quintuplet clusters.

In Figure 3, we show the average time $\langle \tau_B \rangle$ after which the most massive star in a star cluster is a product of binary evolution as a function of the cluster mass M_{cl} for two different primordial binary fractions f_B . The error bars are 1σ standard deviations of 1000 Monte Carlo realizations. The slope of the mass function is $\Gamma = -0.7$, appropriate for the mass-segregated central regions of both Arches and Quintuplet. The more massive a star cluster is, i.e., the more stars that populate the multidimensional binary parameter space, the shorter this average time is because the probability for systems that interact early in their evolution is increased. For less massive clusters, $\langle \tau_B \rangle$ increases, and the statistical uncertainty grows. For example, if $M_{cl} = 10^3 M_\odot$, there are only about 16 ± 3 binaries in which at least one star has a mass above $10 M_\odot$ (for $\Gamma = -0.7$ and $f_B = 100\%$). The same reasoning holds for different binary fractions: the higher the binary fraction is, the more binaries there are, and hence the shorter the average time is until the most massive star results from binary interactions.

With a Salpeter mass function ($\Gamma = -1.35$; Salpeter 1955) the average time until the most massive star is a binary product increases compared to $\Gamma = -0.7$ (Figure 4) because fewer massive binaries interact to form the most massive star. Assume a 4 Myr old star cluster has a mass function slope of $\Gamma = -1.35$, a total mass in stars above $1 M_\odot$ of $M_{cl} = 10^4 M_\odot$ (i.e., a true cluster mass of $1.9 \times 10^4 M_\odot$ if stars below $1 M_\odot$ follow a Kroupa IMF; see Section 2.3), and a primordial binary fraction of $f_B = 60\%$. From Figure 4, we can then read off after what time the most massive star is expected to be a binary product, namely, after 2.5 ± 1.1 Myr.

The central regions of the Arches and Quintuplet clusters have masses of $M_{cl} = 1.5 \times 10^4 M_\odot$ and $0.9 \times 10^4 M_\odot$ in stars

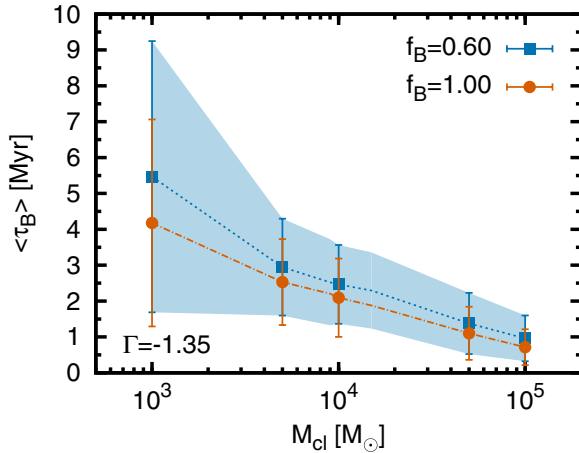


Figure 4. As in Figure 3, but for a steeper mass function with a Salpeter slope of $\Gamma = -1.35$. The binary parameter space spanned by the initial mass ratios and initial orbital separations for massive primary stars is now less populated, resulting in increased average times until the most massive star is a binary product. Similarly, the standard deviations increase.

(A color version of this figure is available in the online journal.)

more massive than $1 M_{\odot}$ (Section 2.3) and ages of 3.5 ± 0.7 and 4.8 ± 1.1 Myr (Section 3.1), respectively. From Figure 3, we expect that the most massive star is a binary product after 1.0 ± 0.7 Myr in the Arches cluster and after 1.7 ± 1.0 Myr in the Quintuplet cluster.

In Figure 5 we show the probability that the most massive star is a binary product and the average number of stars that are more massive than the most massive genuine single star as a function of cluster age for different cluster masses and two different binary fractions. The IMF slope is $\Gamma = -0.7$. The corresponding probabilities and average numbers for a Salpeter ($\Gamma = -1.35$) mass function are shown in Figure 6. Again, the error bars are 1σ standard deviations from 1000 Monte Carlo experiments. Returning to the above-mentioned example star cluster ($M_{cl} = 10^4 M_{\odot}$), from Figure 6, we find that the most massive star is a binary product with a probability of 88% and that the most massive 2.1 ± 1.4 stars are expected to be binary products for the exemplary cluster age of 4 Myr.

Given the ages of the Arches and Quintuplet clusters, we find a probability of $>99.9\%$ that the most massive star in each cluster is a binary product, with the most massive 9.2 ± 3.0 and 7.5 ± 2.8 stars being products of binary evolution in Arches and Quintuplet, respectively. This is compatible with the number of WNh stars in the cores of Arches and Quintuplet, which are the most luminous and hence most massive stars in these clusters, implying they are massive blue stragglers.

5. STELLAR UPPER MASS LIMIT

Data from two star clusters provide the current evidence for the existence of an upper stellar mass limit around $150 M_{\odot}$: the Arches cluster in the Galactic center (Figer 2005) and the R136 cluster in the Large Magellanic Cloud (Weidner & Kroupa 2004; Oey & Clarke 2005; Koen 2006). However, according to our analysis, an upper mass limit cannot be derived from the Arches cluster because (1) it is too old, hence the most massive stars already exploded, and (2) its present-day high-mass star population is dominated by binary products. The situation might be different in the R136 cluster: current age estimates lie in the range 1–4 Myr (Hunter et al. 1995; Massey & Hunter 1998; de Koter et al. 1998; Andersen et al. 2009; Crowther et al. 2010).

In the following we assume that the cluster is young enough that even the most massive stars have not yet evolved off the MS to explore what we can learn from R136 about a possible stellar upper mass limit.

Four stars in R136 with initial masses of $165\text{--}320 M_{\odot}$ appear to exceed the currently discussed upper mass limit of $150 M_{\odot}$ (Crowther et al. 2010). These stars either were born with masses exceeding $150 M_{\odot}$ or gained mass from other stars, e.g., by binary interactions (this work) or dynamically induced stellar mergers (Portegies Zwart et al. 1999; Banerjee et al. 2012b).

From our Monte Carlo simulations (Section 4), we cannot judge with high enough confidence whether the most massive star in R136 is expected to be a binary product or not because of the uncertain age of R136. R136 has an IMF with approximately a Salpeter slope $\Gamma = -1.35$ (Massey & Hunter 1998), and its cluster mass is $5\text{--}10 \times 10^4 M_{\odot}$ (Hunter et al. 1995; Andersen et al. 2009; Hénault-Brunet et al. 2012). From our Monte Carlo simulations of star clusters with binary fractions of 60% and cluster masses M_{cl} of $5 \times 10^4 M_{\odot}$ and $10^5 M_{\odot}$ (Figure 6), we find that the most massive star is expected to be a binary product after 1 Myr with probabilities of 42% and 63%, respectively. The probabilities increase to 74% and 92%, respectively, for a cluster age of 2 Myr and are larger than 98% for an age of 3 Myr. So if the cluster is older than about 2 Myr, the most massive star is likely a binary product (note that our calculations are for a metallicity of $Z = 0.02$, whereas the R136 cluster in the Large Magellanic Cloud has a lower metallicity, so the above numbers will slightly change for the appropriate metallicity but are good enough for this estimate). Because it is not clear whether the most massive star in R136 is a binary product or not, we explore both possibilities.

With Monte Carlo simulations, we investigate the likelihood of finding the observed $280\text{--}320 M_{\odot}$ stars (Crowther et al. 2010) in R136. We randomly sample R136-like star clusters for different adopted stellar upper mass limits M_{up} using the observed IMF slope (Massey & Hunter 1998) of $\Gamma = -1.35$, a binary fraction of 70%, and the fact that R136 contains about 650 stellar systems more massive than $10 M_{\odot}$ (Hunter et al. 1997). We then compute the average number of stars that are initially more massive than a given mass M , $\langle N_{\geq M} \rangle$, and the probability that at least one star is more massive than M , $P_{\geq M}$, by repeating each experiment 1000 times (the quoted errors are 1σ standard deviations). The average numbers and probabilities for the case that binary interactions did not yet take place are summarized in Table 1. For the case where binary interactions already took place, we assume that all massive binaries with initial periods $P_i \leq 5$ days interact by mass transfer (which happens within 2–3 Myr) and that the postinteraction mass is 90% of the total binary mass. The corresponding average numbers and probabilities for this case can be found in Table 2. In both Tables 1 and 2, we also give the results for less massive clusters with 100 and 350 stellar systems initially exceeding $10 M_{\odot}$.

Through binary mergers, stars of up to $300 M_{\odot}$ can be produced if the star formation process stops at an upper mass of $150 M_{\odot}$. However, this scenario requires equal-mass O-type binaries, which are rare (Sana et al. 2012, 2013). We find that with an upper mass of $150 M_{\odot}$, the probability of forming stars in excess of $275 M_{\odot}$ in R136 is zero.¹² With an upper mass limit of $175 M_{\odot}$, the probability of forming at least one star of mass

¹² Given our assumptions, the maximum achievable postinteraction mass is 90% of the total system mass, i.e., $270 M_{\odot}$ for $M_{up} = 150 M_{\odot}$.

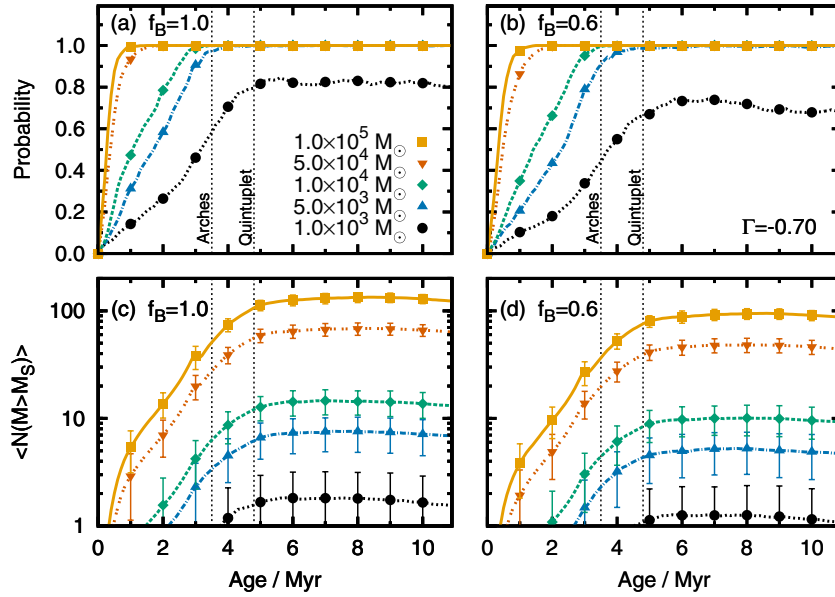


Figure 5. (a) and (b) Probability that the most massive star in a young star cluster is a product of binary evolution and (c) and (d) average number of stars more massive than the most massive genuine single star $\langle N(M > M_S) \rangle$ as a function of age for several cluster masses M_{cl} . The left panels have a binary fraction of 100%, whereas the right panels have a binary fraction of 60%. The symbols represent different cluster masses M_{cl} , and the error bars correspond to the 1σ standard deviation of 1000 realizations per cluster mass. The vertical dashed lines indicate the ages of the Arches and the Quintuplet clusters. The adopted IMF slope is $\Gamma = -0.7$.

(A color version of this figure is available in the online journal.)

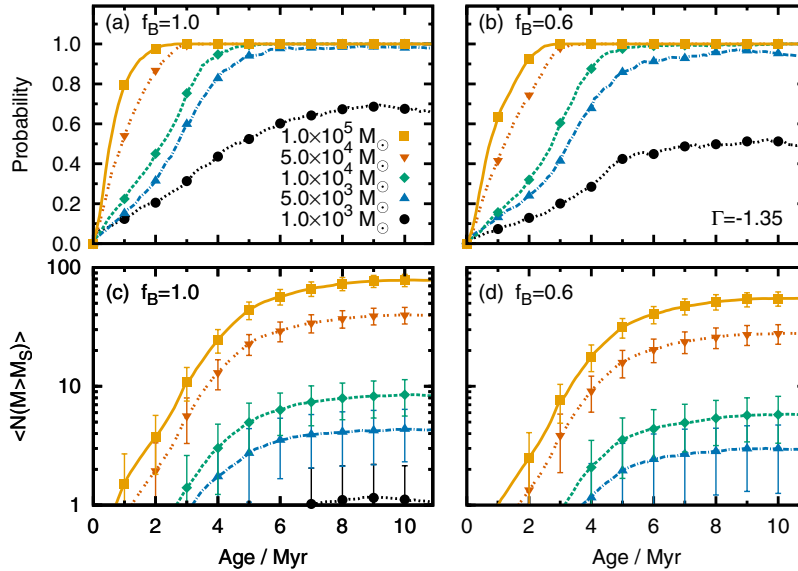


Figure 6. As in Figure 5, but with a Salpeter mass function, i.e., slope of $\Gamma = -1.35$. The relative fraction of massive stars in each cluster is reduced compared to $\Gamma = -0.7$.

(A color version of this figure is available in the online journal.)

$M \geq 275 M_\odot$ increases to 7.0%, and for an upper mass limit of $\sim 200 M_\odot$, the probability of forming at least one star exceeding $275 M_\odot$ and $300 M_\odot$ is 22.8% and 10.6%, respectively. Thus, $200 M_\odot$ provides a lower limit on the maximum stellar birth mass.

It is also unlikely that the upper mass limit exceeds $350 M_\odot$ because then the probability of forming one star above $350 M_\odot$ by binary mass transfer increases to 50.5%, but such massive stars are not observed. We conclude that an upper mass limit in the range of about $200\text{--}350 M_\odot$ is needed to explain the most massive stars in R136 by binary evolution.

Dynamically induced stellar coalescence was proposed as a mechanism to produce the very massive stars in R136

(Portegies Zwart et al. 1999; Banerjee et al. 2012b). However, N -body simulations of dynamically induced stellar coalescence typically produce one to two, in one case up to four, stars exceeding $200 M_\odot$ for R136 when adopting an upper mass limit of $150 M_\odot$ (Banerjee et al. 2012b), i.e., fewer than observed in R136. Furthermore, the rate of dynamically induced mergers in these simulations should be viewed as an upper limit only. Observational results indeed favor a larger half-mass radius (Hunter et al. 1995; Hénault-Brunet et al. 2012) and hence a lower density compared to the simulation assumptions. Similarly, adopting the recent measurements of the orbital distributions of massive binaries (Sana et al. 2012, 2013) further decreases the number of possible dynamical mergers that can

Table 1
Results of Our Monte Carlo Simulations to Determine the Stellar Upper Mass Limit without Binary Interactions

M_{up}/M_{\odot}	M/M_{\odot}	$N_{10} = 100$		$N_{10} = 350$		$N_{10} = 650$	
		$M_{\text{cl}} \approx 2 \times 10^4 M_{\odot}$		$M_{\text{cl}} \approx 7 \times 10^4 M_{\odot}$		$M_{\text{cl}} \approx 10^5 M_{\odot}$	
		$\langle N_{\geq M} \rangle$	$P_{\geq M}$	$\langle N_{\geq M} \rangle$	$P_{\geq M}$	$\langle N_{\geq M} \rangle$	$P_{\geq M}$
10000	150	2.6 ± 1.6	92.0%	9.0 ± 3.0	99.9%	16.9 ± 4.1	>99.9%
	200	1.7 ± 1.3	81.8%	6.1 ± 2.5	99.8%	11.3 ± 3.3	>99.9%
	250	1.3 ± 1.1	71.5%	4.5 ± 2.1	98.2%	8.3 ± 2.9	>99.9%
	300	1.0 ± 1.0	63.4%	3.4 ± 1.8	96.8%	6.5 ± 2.6	99.9%
	350	0.8 ± 0.9	55.1%	2.8 ± 1.6	94.1%	5.2 ± 2.3	99.6%
	400	0.7 ± 0.8	48.5%	2.3 ± 1.5	90.6%	4.4 ± 2.1	99.2%
	450	0.6 ± 0.7	43.0%	2.0 ± 1.4	86.1%	3.7 ± 1.9	98.1%
	500	0.5 ± 0.7	38.5%	1.7 ± 1.3	81.4%	3.2 ± 1.8	96.3%
1000	150	2.5 ± 1.5	92.5%	8.3 ± 2.8	>99.9%	15.4 ± 3.9	>99.9%
	200	1.6 ± 1.2	81.3%	5.4 ± 2.3	99.5%	10.0 ± 3.2	>99.9%
	250	1.1 ± 1.0	69.1%	3.8 ± 2.0	97.5%	7.0 ± 2.7	99.9%
	300	0.8 ± 0.9	58.0%	2.8 ± 1.7	93.2%	5.2 ± 2.3	99.3%
	350	0.6 ± 0.7	48.4%	2.1 ± 1.4	87.8%	4.0 ± 2.1	97.3%
	400	0.5 ± 0.7	40.0%	1.6 ± 1.2	80.3%	3.1 ± 1.8	94.4%
	450	0.4 ± 0.6	32.8%	1.3 ± 1.1	72.6%	2.5 ± 1.6	91.2%
	500	0.3 ± 0.5	27.4%	1.0 ± 1.0	65.0%	2.0 ± 1.4	86.4%
500	150	2.2 ± 1.5	88.5%	7.3 ± 2.6	>99.9%	13.3 ± 3.5	>99.9%
	200	1.3 ± 1.2	73.0%	4.4 ± 2.1	98.5%	8.0 ± 2.7	>99.9%
	250	0.9 ± 0.9	57.4%	2.8 ± 1.7	94.5%	5.1 ± 2.2	99.5%
	300	0.5 ± 0.7	42.5%	1.9 ± 1.3	84.8%	3.3 ± 1.8	96.3%
	350	0.3 ± 0.6	28.7%	1.1 ± 1.0	68.1%	2.0 ± 1.4	87.3%
	400	0.2 ± 0.4	18.0%	0.7 ± 0.8	48.9%	1.1 ± 1.0	68.7%
	450	0.1 ± 0.3	7.9%	0.3 ± 0.5	24.6%	0.5 ± 0.7	39.0%
	500	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
400	150	1.9 ± 1.4	84.2%	6.7 ± 2.5	99.9%	12.6 ± 3.5	>99.9%
	200	1.0 ± 1.0	64.9%	3.7 ± 1.9	97.8%	7.1 ± 2.7	>99.9%
	250	0.6 ± 0.8	43.1%	2.2 ± 1.4	88.6%	4.1 ± 2.1	97.5%
	300	0.3 ± 0.6	26.5%	1.1 ± 1.0	69.2%	2.2 ± 1.5	89.9%
	350	0.1 ± 0.3	11.4%	0.5 ± 0.7	36.3%	0.9 ± 1.0	60.6%
300	400	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	150	1.7 ± 1.3	81.5%	5.6 ± 2.3	99.7%	10.4 ± 3.2	>99.9%
	200	0.8 ± 0.9	53.8%	2.6 ± 1.6	93.1%	4.9 ± 2.2	99.1%
	250	0.3 ± 0.6	26.2%	1.0 ± 1.0	61.4%	1.8 ± 1.3	84.9%
200	300	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	150	0.9 ± 0.9	56.7%	3.0 ± 1.7	95.9%	5.7 ± 2.4	99.8%
	200	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%

Notes. Given are the average number of stars $\langle N_{\geq M} \rangle$ initially more massive than M and the probability $P_{\geq M}$ that at least one star is initially more massive than M for stochastically sampled star clusters as a function of the stellar upper mass limit M_{up} and the number of stars N_{10} more massive than $10 M_{\odot}$. The corresponding total cluster masses M_{cl} extrapolated with a Kroupa IMF (Kroupa 2001) down to $0.08 M_{\odot}$ are also provided. All stars are assumed to be effectively single, i.e., no binary interactions took place.

overcome the $150 M_{\odot}$ limit by a factor of 3.5–4.0. It appears thus unlikely that dynamically induced stellar coalescence is sufficiently efficient to explain the origin of the very massive stars in R136 if the upper mass limit is $150 M_{\odot}$.

As mentioned above, it is also possible that the four massive stars in R136 were born with their deduced initial masses and did not gain mass by other means. This provides then an upper limit on the maximum stellar birth mass. The most massive star found in R136 has an initial mass of $320^{+100}_{-40} M_{\odot}$ (Crowther et al. 2010); hence, the upper mass limit has to be at least of this order, i.e., $\gtrsim 280 M_{\odot}$. This initially $320 M_{\odot}$ star allows us to exclude an upper mass limit of $M_{\text{up}} = 10^4 M_{\odot}$ with 96% confidence because we expect to find 3.2 ± 1.8 stars that initially exceed $500 M_{\odot}$ in this case, although no such star is observed. However, it becomes more difficult to exclude an upper mass limit of $500 M_{\odot}$ or less because the probability of finding *no*

star that initially exceeds $350 M_{\odot}$ ($1 - P_{\geq 350}$) is about 13%; in other words, no star would initially exceed $350 M_{\odot}$ in about every tenth R136-like star cluster for an upper mass limit of $500 M_{\odot}$. The probability increases further to 39% for an upper mass limit of $400 M_{\odot}$. We conclude that stochastic sampling effects are important even in the richest massive star clusters in the Local Group.

Altogether, we find that current data do not exclude an upper mass limit as high as 400 – $500 M_{\odot}$ if binary interactions are neglected. However, the most massive star in R136 is a binary product with a probability of $\gtrsim 40$ – 60% . Including effects of close binary evolution, an initial stellar upper mass limit of at least $200 M_{\odot}$ is required to explain the observed stars with apparent initial masses of about $300 M_{\odot}$. The upper mass limit is thus in the range 200 – $500 M_{\odot}$, thereby solving the maximum mass problem.

Table 2
Results of Our Monte Carlo Simulations to Determine the Stellar Upper Mass Limit with Binary Interactions

M_{up}/M_{\odot}	M/M_{\odot}	$N_{10} = 100$		$N_{10} = 350$		$N_{10} = 650$	
		$M_{\text{cl}} \approx 2 \times 10^4 M_{\odot}$		$M_{\text{cl}} \approx 7 \times 10^4 M_{\odot}$		$M_{\text{cl}} \approx 10^5 M_{\odot}$	
		$\langle N_{\geq M} \rangle$	$P_{\geq M}$	$\langle N_{\geq M} \rangle$	$P_{\geq M}$	$\langle N_{\geq M} \rangle$	$P_{\geq M}$
400	150	2.2 ± 1.5	87.7%	7.6 ± 2.8	99.9%	14.2 ± 3.6	>99.9%
	200	1.3 ± 1.1	70.9%	4.4 ± 2.2	98.4%	8.1 ± 2.8	>99.9%
	250	0.8 ± 0.9	51.4%	2.6 ± 1.6	92.9%	4.8 ± 2.2	99.3%
	300	0.5 ± 0.7	36.1%	1.5 ± 1.2	79.1%	2.8 ± 1.7	93.3%
	350	0.2 ± 0.5	20.8%	0.8 ± 0.9	56.2%	1.5 ± 1.2	75.2%
	400	0.1 ± 0.3	7.1%	0.3 ± 0.5	24.6%	0.5 ± 0.7	39.8%
	450	0.0 ± 0.2	4.7%	0.2 ± 0.5	18.5%	0.3 ± 0.6	26.2%
	500	0.0 ± 0.2	2.7%	0.1 ± 0.3	11.5%	0.2 ± 0.4	16.3%
	550	0.0 ± 0.1	1.6%	0.1 ± 0.3	6.2%	0.1 ± 0.3	9.4%
	600	0.0 ± 0.1	0.7%	0.0 ± 0.2	2.9%	0.0 ± 0.2	4.3%
	650	0.0 ± 0.1	0.3%	0.0 ± 0.1	0.8%	0.0 ± 0.1	1.5%
	700	0.0 ± 0.0	0.1%	0.0 ± 0.0	0.2%	0.0 ± 0.0	0.1%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
350	150	2.1 ± 1.4	87.0%	7.2 ± 2.6	>99.9%	13.6 ± 3.6	>99.9%
	200	1.1 ± 1.1	68.4%	3.9 ± 1.9	98.2%	7.6 ± 2.6	99.9%
	250	0.6 ± 0.8	46.9%	2.1 ± 1.4	88.2%	4.2 ± 2.0	98.8%
	300	0.3 ± 0.5	24.4%	1.1 ± 1.0	66.1%	2.1 ± 1.5	88.0%
	350	0.1 ± 0.3	8.3%	0.3 ± 0.6	28.2%	0.7 ± 0.8	50.5%
	400	0.0 ± 0.2	4.2%	0.2 ± 0.4	16.4%	0.4 ± 0.6	33.0%
	450	0.0 ± 0.1	2.1%	0.1 ± 0.3	8.3%	0.2 ± 0.5	20.7%
	500	0.0 ± 0.1	1.1%	0.0 ± 0.2	3.2%	0.1 ± 0.3	10.4%
	550	0.0 ± 0.1	0.3%	0.0 ± 0.1	1.6%	0.0 ± 0.2	2.9%
	600	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	650	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	700	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
300	150	1.9 ± 1.4	86.5%	6.5 ± 2.5	>99.9%	12.3 ± 3.5	>99.9%
	200	1.0 ± 0.9	63.8%	3.3 ± 1.7	97.5%	6.1 ± 2.5	99.8%
	250	0.5 ± 0.7	38.4%	1.5 ± 1.2	79.1%	2.8 ± 1.7	93.0%
	300	0.1 ± 0.4	12.7%	0.4 ± 0.6	33.7%	0.7 ± 0.8	51.5%
	350	0.1 ± 0.3	7.7%	0.2 ± 0.5	19.1%	0.4 ± 0.6	30.4%
	400	0.0 ± 0.2	3.7%	0.1 ± 0.3	10.1%	0.2 ± 0.4	15.4%
	450	0.0 ± 0.1	1.5%	0.0 ± 0.2	3.3%	0.1 ± 0.2	5.5%
	500	0.0 ± 0.1	0.4%	0.0 ± 0.1	0.7%	0.0 ± 0.1	0.8%
	550	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	600	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	650	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	700	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
250	150	1.6 ± 1.3	78.2%	5.5 ± 2.5	99.8%	10.6 ± 3.3	>99.9%
	200	0.7 ± 0.8	47.2%	2.4 ± 1.6	89.8%	4.4 ± 2.1	98.5%
	250	0.1 ± 0.4	12.7%	0.5 ± 0.7	39.6%	1.0 ± 1.0	61.6%
	300	0.1 ± 0.2	4.8%	0.2 ± 0.5	20.5%	0.4 ± 0.7	36.1%
	350	0.0 ± 0.1	1.6%	0.1 ± 0.3	7.9%	0.2 ± 0.4	15.4%
	400	0.0 ± 0.1	0.3%	0.0 ± 0.1	1.7%	0.0 ± 0.2	2.8%
	450	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	500	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	550	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	600	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	650	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	700	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
200	150	1.2 ± 1.1	69.2%	4.1 ± 2.0	98.4%	7.5 ± 2.8	>99.9%
	200	0.2 ± 0.4	20.1%	0.8 ± 0.9	52.9%	1.4 ± 1.1	76.3%
	250	0.1 ± 0.3	8.0%	0.3 ± 0.5	23.7%	0.5 ± 0.7	38.5%
	300	0.0 ± 0.1	1.8%	0.1 ± 0.3	6.7%	0.1 ± 0.4	10.6%
	350	0.0 ± 0.0	0.1%	0.0 ± 0.0	0.2%	0.0 ± 0.0	0.2%
	400	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	450	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	500	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	550	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	600	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	650	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	700	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
150	150	0.3 ± 0.5	23.6%	1.1 ± 1.0	66.7%	2.0 ± 1.4	87.8%
	200	0.1 ± 0.3	7.2%	0.3 ± 0.5	24.9%	0.5 ± 0.7	39.6%
	250	0.0 ± 0.1	0.6%	0.0 ± 0.1	1.7%	0.0 ± 0.2	3.5%
	300	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	350	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	400	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	450	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	500	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	550	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	600	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	650	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	700	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	750	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%
	800	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%	0.0 ± 0.0	0.0%

Notes. As in Table 1, but now it is assumed that binary interactions took place in all massive binaries with initial orbital periods shorter than 5 days such that masses higher than the stellar upper mass limit can be achieved because of binary mass transfer. Stars with masses smaller than the upper mass limit are either (effectively) single stars or again products of binary mass exchange in binaries with initial orbital periods shorter than 5 days.

6. UNCERTAINTIES

There are several sources of uncertainty that affect theoretical and observed mass functions and hence, e.g., our cluster ages derived from them. It is important to understand the uncertainties to estimate their influence on our conclusions and the derived quantities. The conclusion that binary effects shape the upper end of the stellar mass function remains unaffected. In Section 6.1, we discuss modeling uncertainties because of the fitting procedure, stellar wind mass loss, binary star evolution, and rotation. Observational uncertainties such as the influence of different reddening laws on derived stellar masses of stars in the Galactic center are discussed in Section 6.2. We discuss the influence of dynamical interactions on stellar mass functions in Section 6.3. Star formation histories that are different from single starbursts are considered in Section 6.4 to investigate whether such scenarios are also consistent with the observed age spread among the most massive stars and the resulting stellar mass functions.

6.1. Modeling Uncertainties

6.1.1. Fitting Uncertainties

Stars in the wind mass loss peak of the mass function will very soon leave the MS. The mass of these turn-off stars is a sensitive function of cluster age, especially for massive stars, which radiate close to the Eddington limit. Massive stars have lifetimes that depend only weakly on mass, and hence a small change in age corresponds to a large change in mass. We cannot reproduce the observed mass functions of Arches and Quintuplet if we change the age of our models in Figure 2 by more than 0.2–0.3 Myr. We therefore adopt 0.3 Myr as the age uncertainty associated with our fitting.

The initial binary fraction is best constrained by the number of stars in the mass function tail because it consists only of either postinteraction or preinteraction unresolved binaries. In contrast, the wind mass loss peak changes little with the binary fraction. Our observational sample is limited by the exclusion of WNh stars in both Arches and Quintuplet because no reliable masses of the WNh stars could be determined (see Section 2.4). We thus cannot determine the primordial binary fractions accurately, especially because the tail of the Quintuplet mass function is not very pronounced. Increasing the age of our Quintuplet model by 0.1 Myr allows for 100% binaries while maintaining a satisfactory, albeit slightly inferior to the best, fit to the mass function. Both clusters are thus consistent with having the same primordial binary fraction.

6.1.2. Stellar Wind Mass Loss

Our wind mass loss prescription (Nieuwenhuijzen & de Jager 1990) slightly underestimates stellar winds compared to the latest predictions (Vink et al. 2000, 2001). Compared to the most recent stellar evolution models of Ekström et al. (2012), which use the prescriptions of Vink et al. (2000, 2001), we find that our turn-off masses agree to within 2%–3% for initial masses $\lesssim 50 M_{\odot}$, while in more massive stars our turn-off mass is up to 15% larger, mainly because of the applied Wolf–Rayet wind mass loss rates in Ekström et al. (2012).

The widths of the bins in our model mass functions are 0.04 dex; that is, masses differ by about 10% from bin to bin. The observed mass functions have bin sizes of 0.2 dex; that is, masses are different by 59% from bin to bin. Wind mass loss prescriptions that lead to stellar masses at the end of the MS

that differ by only a few percent result in indistinguishable mass functions; our mass functions and conclusions are therefore essentially independent of whether the empirical wind mass loss prescription of Nieuwenhuijzen & de Jager (1990) or the theoretical prescription of Vink et al. (2000, 2001) is used.

Augmenting our wind loss rate by 70%, we find that an initially $85 M_{\odot}$ star has a turn-off mass of about $49 M_{\odot}$, which matches the recent stellar models by Ekström et al. (2012) (compared to $\sim 58 M_{\odot}$ in our standard model). With the enhanced wind mass loss rate, the Arches wind mass loss peak corresponds to an initially $\sim 85 M_{\odot}$ star with a MS lifetime of 3.3 Myr, compared to $70 M_{\odot}$ and 3.5 Myr, respectively, in our standard model. The Quintuplet wind mass loss peak comes from initially $\sim 40 M_{\odot}$ stars, for which the uncertainty in wind mass loss is $< 3\%$. Our Quintuplet age estimate is thus robust with respect to the wind mass loss rate uncertainty.

6.1.3. Binary Star Evolution

Our understanding of binary star evolution in general is subject to uncertainties. Uncertainties that directly influence the shape of the mass function tails are discussed in Schneider et al. (2013). A further, more quantitative discussion of uncertainties in binary star evolution is found in de Mink et al. (2013a, 2013b). Here, we restrict ourselves to MS stars, i.e., to mergers of two MS stars and mass transfer onto MS stars. Mergers that involve a post-MS star form a post-MS object and are thus not considered here.

We assume that two MS stars merge if the mass ratio of the accretor to donor star is less than 0.56. This threshold is calibrated against the detailed binary models of de Mink et al. (2007) and is of limited relevance to our results: if a binary does not merge but instead transfers mass (or vice versa), the accretor becomes massive because the mass transfer efficiency of MS stars is high (e.g., Wellstein et al. 2001; Langer 2012). In either case, the mass gainer will be a massive star (de Mink et al. 2013a). The expected binary fraction of stars in the tail of the mass functions, however, changes: a lower critical mass ratio leads to fewer MS mergers and hence to a higher binary fraction and vice versa.

The amount of rejuvenation of MS mergers is determined by the amount of mixing of fresh fuel into the core of the merger product and determines by how much the lifetime of the merger product is prolonged. The more rejuvenation there is, the longer the remaining MS lifetime is and the more mergers are expected to be found. We assume that a fraction of 10% of the envelope is mixed into the core, resulting into fairly short remaining MS lifetimes of the merger products compared to the assumption of complete mixing used in the original Hurley et al. (2002) code. Recent simulations of massive mergers seem to support the mild mixing as used in our work (Glebbeek et al. 2013, and references therein).

The mass transfer efficiency is important for our results. The more transferred mass is accreted during RLOF, the larger the final mass of the accreting star is. The maximum reachable mass of any accretor is given by the total mass of the binary (i.e., at most twice the mass of the donor star), and the larger the overall mass transfer efficiency is, the more binary products exceed the most massive genuine single star. In our models, we limit the mass accretion rate to the thermal timescale of the accretor, which results in higher mass transfer efficiencies the larger the mass ratio is and the closer the binary is (see Schneider et al. 2013). This idea is motivated by detailed binary evolution models (e.g., Ulrich & Burger 1976; Kippenhahn &

Meyer-Hofmeister 1977; Neo et al. 1977; Pols & Marinus 1994; Wellstein et al. 2001).

The initial distributions of primary masses, mass ratios, and orbital separations determine the relative fraction of stars that will merge, transfer mass, etc. It turns out that the distribution of orbital separations influences the incidence of binary products most (de Mink et al. 2013a, 2013b) because initially close binaries transfer mass, on average, more efficiently than wider binaries. A distribution of initial orbital separations that favors close binaries therefore leads to, on average, more massive binary products than distribution functions that favor initially wider binaries.

A more quantitative assessment of the above-quoted uncertainties in binary evolution and initial binary distribution functions reveals that a population of MS stars with luminosities $L > 10^4 L_\odot$ (i.e., O and B stars) contains about $30^{+10}_{-15}\%$ binary products if continuous star formation is assumed (de Mink et al. 2013b). All in all, depending on the exact assumptions regarding binary evolution, there will be more or fewer stars in the tail of the mass function. However, the peak–tail structure never disappears unless it is assumed that neither MS mergers nor RLOF are able to increase stellar masses, which is unphysical.

6.1.4. Stellar Rotation

Mixing induced by stellar rotation increases the fuel available to a star and increases its lifetime. The amount of mixing grows with increasing mass, increasing rotation rate, and decreasing metallicity and may contribute to the observed age spreads and mass function tails in Arches and Quintuplet. The models of Brott et al. (2011) show that the MS lifetime of a $60 M_\odot$ star lengthens by 0.2 and 0.6 Myr for initial rotational velocities of 300 and 500 km s^{-1} , respectively. Assuming that the present-day distribution of rotational velocities of Galactic O and B stars approximately represents the initial distribution, no more than 10% and less than 1% of stars would have initial rotation rates exceeding 300 and 500 km s^{-1} , respectively, and are thus expected to be influenced significantly by rotational mixing (see Table 2 in de Mink et al. 2013a and references therein). This is small compared to the 40% of all O stars that undergo strong binary interaction during their MS evolution (Sana et al. 2012). The present-day distribution of rotational velocities is probably altered, e.g., by binary star evolution such that some of the fast rotators are expected to have gained their fast rotation by binary interactions (de Mink et al. 2013, but see also Ramírez-Agudelo et al. 2013). In this respect, the expected fraction of genuine Galactic single stars that are significantly affected by rotational mixing is even smaller than the above-quoted fractions. The effect of rotation is thus only of limited relevance to our results compared to binary star interactions.

6.2. Observational Uncertainties

There are two steps involved in determining stellar masses from photometric observations that contribute to the uncertainties of the derived individual stellar masses. The first step involves the conversion of the observed apparent magnitudes (fluxes) to absolute magnitudes and luminosities, respectively, taking, among other things, the distance and extinction into account. The second step involves the conversion of luminosities to stellar masses. This step relies upon mass–luminosity relations that depend on (in general, a priori unknown) stellar ages and the applied stellar models. In this section, we estimate the uncertainty on individual stellar masses introduced by these two

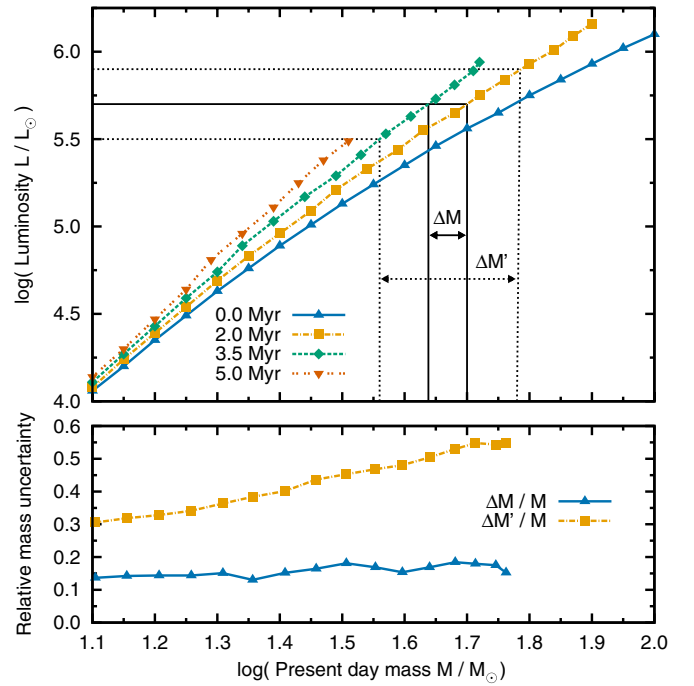


Figure 7. Uncertainties on derived individual stellar masses introduced by unknown a priori stellar ages, $\Delta M / M$, and by the combined effect of unknown stellar ages and uncertain luminosities, $\Delta M' / M$. The top panel illustrates how we estimate these uncertainties for a star with luminosity $\log L / L_\odot = 5.7 \pm 0.2$ using the MS mass–luminosity relations of different ages from our code (Hurley et al. 2000). The bottom panel shows the uncertainties as a function of mass.

(A color version of this figure is available in the online journal.)

steps for stars in Arches and Quintuplet. Once we know the uncertainties, we can apply them to the turn-off masses derived from the wind mass loss peak in the mass function to find the corresponding uncertainty in cluster age.

In the top panel of Figure 7 we show MS mass–luminosity relations of Milky Way stars of different ages as used in our code (based on Hurley et al. 2000). Not knowing the exact age of a star, but only a probable age range (here 2.0–3.5 Myr), introduces an uncertainty ΔM on the derived individual stellar masses (see bottom panel in Figure 7). The more evolved and the more massive a star is, the larger the uncertainty is. If we additionally include the uncertainty in the luminosity (here ± 0.2 dex; see below), the uncertainty in the stellar mass grows to $\Delta M'$. The uncertainty in luminosity is the dominant contribution here.

The uncertainties presented in Figure 7 are tailored to stars in the Arches cluster where stellar masses have been derived by Stolte et al. (2005) using a 2 Myr isochrone (mass–luminosity relation), whereas our analysis reveals an age of about 3.5 Myr, hence the 2.0–3.5 Myr age range for the mass–luminosity relations. Deriving individual stellar masses using two different extinction laws, Habibi et al. (2013) find that stellar masses can deviate by up to 30% when either a Nishiyama et al. (2009) extinction law or the traditional Rieke & Lebofsky (1985) law toward the Galactic center is used. As the appropriate extinction law toward the Galactic center line of sight is still a matter of debate, we adopt uncertainties of ± 0.2 dex on luminosities to be in line with the work of Habibi et al. (2013). The situation is similar in Quintuplet: we find slightly smaller uncertainties on derived stellar masses than given in Figure 7 for uncertainties of ± 0.2 dex on luminosities and an age range of 4–5 Myr. To be conservative, we adopt an uncertainty of $\pm 30\%$ (i.e.,

$\Delta M'/M = 60\%$) for individual stellar masses in both clusters, which agrees with the diversity of mass estimates in the literature for stars in Arches.

Habibi et al. (2013) find stellar masses that are up to 30% less than those of our analysis for a Nishiyama et al. (2009) extinction law. An independent study of the Arches present-day mass function (Kim et al. 2006) shows a similar peak–tail structure at the high-mass end using a different extinction law, and the derived masses are comparable to those of our analysis (Stolte et al. 2005). Masses in excess of $150 M_{\odot}$, i.e., larger than our masses, have been suggested for the most luminous stars in the Arches cluster by Crowther et al. (2010). Hence, an uncertainty of $\pm 30\%$ applied to our adopted stellar masses covers the complete range of suggested masses for stars in Arches.

The overall structure of the mass functions of Arches and Quintuplet—a stellar wind peak and binary tail—is robust to the mentioned uncertainties. We also find the peak–tail structure in the mass function of Habibi et al. (2013) and in mass functions constructed from the photometric data of Stolte et al. (2005) and Hußmann et al. (2012), respectively, using isochrones of different ages. The whole mass functions shift in mass, but the relative structure stays the same. The reason is that different extinction laws or isochrones systematically influence all stars in a similar way and do not introduce differential effects.

Applying $\pm 30\%$ uncertainties on stellar masses leads to turn-off masses of the Arches and Quintuplet of about $70 \pm 21 M_{\odot}$ and $40 \pm 12 M_{\odot}$, with associated age uncertainties of ± 0.6 and ± 1.1 Myr, respectively.

6.3. Dynamical Interactions in Star Clusters

The observed present-day mass functions are influenced by dynamical cluster evolution (Harfst et al. 2010 and references therein). Flat mass function slopes of the order of $\Gamma = -0.7$, compared to a Salpeter IMF slope of $\Gamma = -1.35$, are likely a consequence of mass segregation in which massive stars sink toward the cluster center. For the Arches cluster, Habibi et al. (2013) show that a dynamical model with a standard Salpeter IMF explains the steepening of the IMF slope toward larger distances from the cluster center. While the conclusion is more elusive for the dispersed Quintuplet population, the similarity of the mass function slopes in the inner regions of both clusters and the older age of Quintuplet suggest that similar processes have shaped the present-day mass function of this cluster as well.

We further investigate whether dynamical interactions in star clusters can also cause the peak in the observed mass functions. Mass segregation flattens the high-mass end of the mass function without producing a peak (Portegies Zwart et al. 2007), while dynamical ejection of stars works on all stars with an ejection efficiency monotonically increasing with mass (Fujii & Portegies Zwart 2011; Perets & Šubr 2012; Banerjee et al. 2012a), as confirmed by the observed Galactic fraction of runaway O stars, which is larger than that of runaway B stars (Gies & Bolton 1986; Stone 1991). Such a smoothly increasing ejection efficiency does not create a peak in the mass function but can give rise to a tail that is, however, not seen in the mass functions of N -body simulations of the Arches cluster (Portegies Zwart et al. 2007; Harfst et al. 2010).

6.4. Star Formation Histories

Star formation is not an instantaneous process but lasts a finite amount of time. Observations show that there is an empirical

relationship between the duration of star formation and the crossing time of star clusters (Elmegreen 2000). For example, in Westerlund 1 and NGC 3603YC the age difference among stars less massive than about $11.5 M_{\odot}$ and $6.5 M_{\odot}$ shows that they were formed within at most 0.4 and 0.1 Myr (Kudryavtseva et al. 2012), respectively, which compares well to the cluster crossing times of 0.3 Myr (Brandner et al. 2008) and 0.1 Myr (Pang et al. 2010).

The core of the Arches cluster has a radius of about 0.23 pc (Figer et al. 2002) and a velocity dispersion of 5.7 km s^{-1} (Clarkson et al. 2012) at a distance of 8.0 kpc, corresponding to a crossing time, and hence a duration of star formation, of about 0.04 Myr. A similar estimate for the Quintuplet cluster is more uncertain because the core radius and especially the velocity dispersion are not well known. If we assume that the observed central region of Quintuplet with a radius of about 0.5 pc (Hußmann et al. 2012) corresponds to the core radius and that the velocity dispersion is about 17 km s^{-1} (Liermann et al. 2009), the crossing time is 0.03 Myr. The mass functions produced by such short periods of star formation are indistinguishable from an instantaneous starburst. The apparent age spread among the most luminous stars in Arches and Quintuplet is about 1–1.5 Myr (Martins et al. 2008; Liermann et al. 2012) and hence much larger than the estimated star formation periods.

In the Appendix we investigate whether a star formation history that is different from a single starburst can also explain the observed peak–tail structure in the mass functions of Arches and Quintuplet. We find that this is possible with, e.g., a two-stage starburst, but the age spread among the most massive stars would then be inconsistent with observations.

7. CONCLUSIONS

Massive stars rapidly change their mass, thereby altering the stellar mass function. Stellar wind mass loss reduces stellar masses such that stars accumulate near the high-mass end of present-day mass functions, creating a bump whose position reveals the mass of the turn-off stars and hence the age of young star clusters (Schneider et al. 2013). Binary stars are frequent and important for massive star evolution because of binary mass exchange (Sana et al. 2012): mass transfer and stellar mergers increase stellar masses and create a tail of rejuvenated binary products at the high-mass end of mass functions (Schneider et al. 2013). We model the observed mass functions of the young Arches and Quintuplet star clusters (Stolte et al. 2005; Hußmann et al. 2012) using a rapid binary evolution code (Hurley et al. 2002; Izzard et al. 2004, 2006, 2009) to identify these two features and to address two pressing controversies.

1. *The cluster age problem.* The most massive stars in Arches and Quintuplet, the WNh stars, appear to be younger than the less massive O stars (Martins et al. 2008; Liermann et al. 2012). This is not expected from star cluster formation (Elmegreen 2000; Kudryavtseva et al. 2012) but is well known in older clusters under the blue straggler phenomenon.
2. *The maximum mass problem.* A stellar upper mass limit of $150 M_{\odot}$ is observationally determined (Weidner & Kroupa 2004; Figer 2005; Oey & Clarke 2005; Koen 2006). In contrast, SN 2007bi is thought to be a pair-instability supernova from an initially $250 M_{\odot}$ star (Gal-Yam et al. 2009; Langer 2009), and four stars greatly exceeding this limit are found in the R136 cluster in the Large Magellanic Cloud (Crowther et al. 2010).

We identify the peak and tail in the observed mass functions of the Arches and Quintuplet clusters. By fitting our models to the observations, we determine the ages of Arches and Quintuplet to be 3.5 ± 0.7 and 4.8 ± 1.1 Myr, respectively. This solves the age problem because the most massive stars in Arches and Quintuplet are rejuvenated products of binary mass transfer, and the ages derived from these stars are therefore significantly underestimated. While our age error bars are still large, mostly because of uncertain absolute magnitudes, our method removes the ambiguity in the age determination. For the age determination of older star clusters, blue stragglers are eminently disregarded when isochrones are fitted to the turn-off in Hertzsprung–Russell diagrams. Our analysis shows that for young star clusters, where the higher fraction of interacting binaries produces even more blue stragglers, they obviously need to be disregarded as well in order to derive the correct cluster age.

Even without modeling the observed mass functions, the ages and also the mass lost by the turn-off stars during their MS evolution can be determined from the mass function alone. Refilling the mass function above the present-day mass of the turn-off stars with the number of excess stars in the wind mass loss peak gives the initial mass of the turn-off stars and hence the cluster age. The derived ages agree with the more accurate ages from detailed modeling of the observed mass functions. According to this new method, the turn-off stars lost $12\text{--}17 M_{\odot}$ of their initial $62\text{--}72 M_{\odot}$ in Arches and $4\text{--}8 M_{\odot}$ of their initial $36\text{--}47 M_{\odot}$ in Quintuplet. This is the first direct measurement of stellar wind mass loss that does not rely on derivations of stellar wind mass loss rates.

Monte Carlo experiments to investigate the effects of stochastic sampling show that the most massive star in the Arches and Quintuplet clusters is expected to be a rejuvenated product of binary mass transfer after 1.0 ± 0.7 and 1.7 ± 1.0 Myr, respectively. At their present age, the probability that the most massive star in Arches and Quintuplet is a product of binary mass exchange is $>99.9\%$, and the most massive 9.2 ± 3.0 and 7.5 ± 2.8 stars in Arches and Quintuplet, respectively, are expected to be such rejuvenated binary products.

Our findings have implications for the maximum mass problem. The Arches cluster is older than previously thought, and its most massive stars are most likely binary products. The mass function is thus truncated by finite stellar lifetimes and not by an upper mass limit. To constrain a potential stellar upper mass limit, we consider the massive cluster R136 in the Large Magellanic Cloud, which is thought to be so young that its initially most massive stars are still alive today. We find that the most massive star is a binary product with a probability of $>40\%$, depending on the exact, albeit uncertain, cluster age (Section 5). Assuming binaries already interacted, a stellar upper mass limit of at least $200 M_{\odot}$ is needed to form the observed $165\text{--}320 M_{\odot}$ stars in R136. It can also not exceed $350 M_{\odot}$ because then the probability of forming stars above, e.g., $350 M_{\odot}$ becomes larger than about 50% , but such stars are not observed. Assuming that no binary interactions changed the masses of the very massive stars in R136, a stellar upper mass limit of up to $400\text{--}500 M_{\odot}$ cannot be fully excluded because of stochastic sampling even in this rich star cluster. The upper mass limit is thus likely in the range $200\text{--}500 M_{\odot}$, thereby solving the maximum mass problem.

We conclude that the most massive stars in the universe may be the rejuvenated products of binary mass transfer. Because of their extreme mass and luminosity, radiation feedback from

these stars is crucial to observable properties of young stellar populations, to the state of the interstellar medium around young stellar clusters, and even to the reionization of the universe after the big bang. Our results have strong implications for understanding star-forming regions nearby and at high redshift as observationally derived fundamental properties like IMFs are based on the assumption that the brightest stars are single and less massive than $150 M_{\odot}$. These very massive stars are thought to die as pair-instability supernovae and produce huge, so far unaccounted, contributions to the chemical enrichment of nearby and distant galaxies (Langer 2012), and their final explosions may be observable throughout the universe. Understanding the most massive stars in young nearby star clusters is an essential step toward investigating these exciting phenomena that shape our cosmos.

We thank the referee, Dany Vanbeveren, for carefully reading our manuscript and constructive suggestions. F.R.N.S. acknowledges the fellowships awarded by the German National Academic Foundation (Studienstiftung) and the Bonn-Cologne Graduate School of Physics and Astronomy. R.G.I. thanks the Alexander von Humboldt foundation. S.d.M. acknowledges support by NASA through Hubble Fellowship grant HST-HF-51270.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA under contract NAS 5-26555, and the Einstein Fellowship program through grant PF3-140105 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under contract NAS8-03060. B.H. and A.S. acknowledge funding from the German science foundation (DFG) Emmy Noether program under grant STO 496-3/1.

APPENDIX

STAR FORMATION HISTORIES CONTINUED

Here we investigate changes in the mass function due to a star formation scenario that deviates from a true starburst in order to understand whether the observed mass functions of Arches and Quintuplet can be reproduced without binaries. We analyze two scenarios: (1) a period of prolonged but constant star formation rate and (2) two instantaneous starbursts separated in time. The latter scenario represents not only a two-stage starburst within one cluster but also two merged star clusters where stars in each cluster are coeval. This situation most probably applies to the massive star cluster R136 in the Large Magellanic Cloud, which is thought to be a double cluster in the process of merging (Sabbi et al. 2012). We compute mass functions for star formation scenarios 1 and 2 (which include single, true starbursts) and search for parameter values that minimize the least squares deviation of the modeled ($y_{\text{model},i}$) from the observed ($y_{\text{obs},i}$) mass functions of the Arches and Quintuplet clusters assuming Poisson uncertainties, i.e., $\sigma_{\text{obs},i}^2 = y_{\text{obs},i}$,

$$\chi^2 = \frac{1}{N} \sum_{i=1}^N \frac{(y_{\text{model},i} - y_{\text{obs},i})^2}{\sigma_{\text{obs},i}^2}, \quad (\text{A1})$$

where N is the number of mass bins. Exemplary star formation scenarios are described in Table 3, and the resulting mass functions are compared to observations in Figure 8. Among these examples are those star formation scenarios that lead to the best fits (models A2, A4, Q2, and Q4).

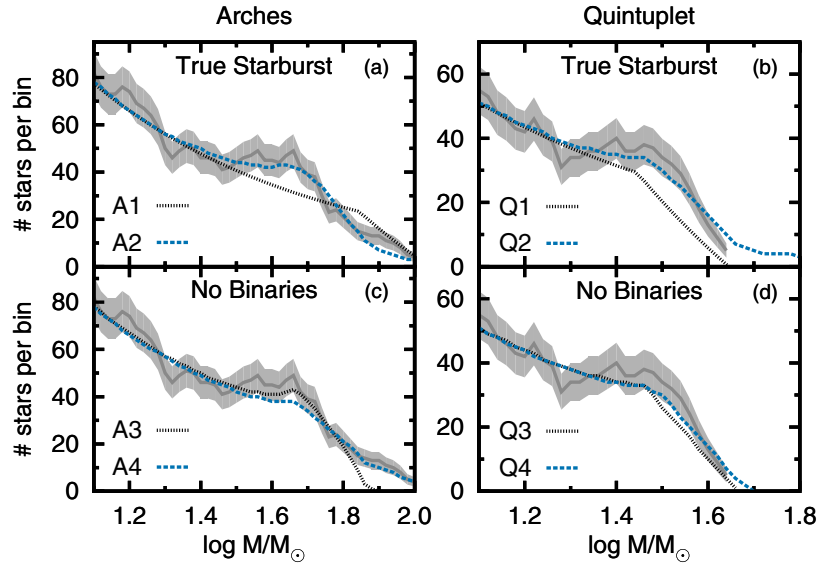


Figure 8. Modeled mass functions using different star formation scenarios compared to observations of the (a) and (c) Arches and (b) and (d) Quintuplet clusters. In the top panels, we compare the observations to our best-fitting starburst models including binary stars and to simple power-law mass functions truncated at the observed maximum mass. In the bottom panels, we show mass functions composed only of single stars with more complex star formation scenarios. The individual star formation models A1–A4 and Q1–Q4 are explained in Table 3 together with their least squares deviation χ^2 (see Equation(A1)). All modeled mass functions are binned in the same way as the observations (see Section 2.5).

(A color version of this figure is available in the online journal.)

Table 3
Star Formation (SF) Scenarios Used to Compute the Mass Functions that Are Compared to Observations in Figure 8

SF Model	t_1/Myr	t_2/Myr	f_B	χ^2_{tot}	χ^2_{peak}	Description
A1	1.19	1.59	Power-law mass function truncated at the most massive observed star; power-law index $\Gamma = -0.7$
A2	3.5	...	100%	0.57	0.36	Single starburst at t_1
A3	3.2	3.3	0%	1.85	0.31	Constant SF between t_1 and t_2
A4	0.7	3.3	0%	0.49	0.69	Two starbursts at t_1 and t_2
Q1	2.11	4.23	Power-law mass function truncated at the most massive observed star; power-law index $\Gamma = -0.7$
Q2	4.8	...	60%	0.56	0.36	Single starburst at t_1
Q3	4.8	...	0%	0.70	0.97	Single starburst at t_1
Q4	4.5	4.7	0%	0.43	0.46	Constant SF between t_1 and t_2

Notes. Given are the primordial binary fraction f_B for each model as well as the least squares deviation χ^2_{tot} in the total mass range ($1.1 \leq \log M/M_\odot \leq 2.0$) and χ^2_{peak} in a mass region around the wind mass loss peak ($1.4 \leq \log M/M_\odot \leq 1.8$ for Arches and $1.3 \leq \log M/M_\odot \leq 1.6$ for Quintuplet). The best-fit models are A2, A4, Q2, and Q4.

From Table 3 and the top panels of Figure 8 it is evident that simple power-law mass functions (models A1 and Q1) fit the observed mass functions of Arches and Quintuplet much worse than the best single-starburst models including binary stars (A2 and Q2). Particularly, the mass region around the wind mass loss peak is not fitted well by models A1 and Q1 (see χ^2_{peak} in Table 3). The mass functions of the Arches and Quintuplet clusters do not follow simple power laws.

In the bottom panels of Figure 8, we also present models of the observed mass functions of the Arches and Quintuplet clusters *without* binary stars. We do not find satisfactory models that fit peak and tail simultaneously with a single starburst without binaries. Model A3, for example, fits the peak due to mass loss well ($\chi^2_{\text{peak}} = 0.31$) but fails to explain the high-mass end of the observed mass function ($\chi^2_{\text{tot}} = 1.85$).

We can improve this situation by adding a younger stellar population that fits the tail. Such a scenario is given by model A4, which fits the peak and the total mass function. However, this two-component model has an age spread of 2.6 Myr, which is more than twice as large as the observed age discrepancy of about 1 Myr in the Arches cluster (Martins

et al. 2008) and much larger than the expected star formation duration.

A two-component solution is not needed to model the observed mass function of the Quintuplet cluster because the tail of the mass function is not very pronounced. Consequently, our models Q3 and Q4 predict no or a too small age spread, contrary to the observations.

The age spread of 0.2 Myr of model Q4 might be compatible with the above-estimated star formation duration given the quite uncertain core radius and velocity dispersion of Quintuplet. The single-starburst model, Q3, is shown to illustrate the difference between the mass functions with (Q2) and without (Q3) binaries. The tail of the mass function is, however, underestimated in the observed mass function in Figure 8 because no self-consistent mass determination for the three WNH stars in the core of Quintuplet is available (Hußmann et al. 2012). If the tail were visible, we could, of course, model it by an additional younger stellar population as is done in Arches.

In summary, we conclude that we can reproduce the mass functions of Arches and Quintuplet without binaries but with freedom in the star formation history. However, the best-fit

star formation parameters (e.g., the age spread of 2.6 Myr) are inconsistent with other observables. Our single-starburst models that include binaries are thus the only models that fulfill all observational constraints. These models are also consistent with a star formation duration of the order of the crossing time of the cluster.

REFERENCES

- Andersen, M., Zinnecker, H., Moneti, A., et al. 2009, *ApJ*, **707**, 1347
- Banerjee, S., Kroupa, P., & Oh, S. 2012a, *ApJ*, **746**, 15
- Banerjee, S., Kroupa, P., & Oh, S. 2012b, *MNRAS*, **426**, 1416
- Blum, R. D., Schaerer, D., Pasquali, A., et al. 2001, *AJ*, **122**, 1875
- Brandner, W., Clark, J. S., Stolte, A., et al. 2008, *A&A*, **478**, 137
- Brott, I., de Mink, S. E., Cantiello, M., et al. 2011, *A&A*, **530**, A115
- Chen, X., & Han, Z. 2009, *MNRAS*, **395**, 1822
- Clarkson, W. I., Ghez, A. M., Morris, M. R., et al. 2012, *ApJ*, **751**, 132
- Collier, A. C., & Jenkins, C. R. 1984, *MNRAS*, **211**, 391
- Crowther, P. A., Schnurr, O., Hirschi, R., et al. 2010, *MNRAS*, **408**, 731
- de Koter, A., Heap, S. R., & Hubeny, I. 1998, *ApJ*, **509**, 879
- de Mink, S. E., Langer, N., Izzard, R. G., Sana, H., & de Koter, A. 2013a, *ApJ*, **764**, 166
- de Mink, S. E., Pols, O. R., & Hilditch, R. W. 2007, *A&A*, **467**, 1181
- de Mink, S. E., Sana, H., Langer, N., Izzard, R. G., & Schneider, F. R. N. 2013b, *ApJ*, submitted
- Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, *A&A*, **537**, A146
- Elmegreen, B. G. 2000, *ApJ*, **530**, 277
- Espinoza, P., Selman, F. J., & Melnick, J. 2009, *A&A*, **501**, 563
- Figer, D. F. 2005, *Natur*, **434**, 192
- Figer, D. F., McLean, I. S., & Morris, M. 1999, *ApJ*, **514**, 202
- Figer, D. F., Najarro, F., Gilmore, D., et al. 2002, *ApJ*, **581**, 258
- Figer, D. F., Najarro, F., Morris, M., et al. 1998, *ApJ*, **506**, 384
- Fujii, M. S., & Portegies Zwart, S. 2011, *Sci*, **334**, 1380
- Gal-Yam, A., Mazzali, P., Ofek, E. O., et al. 2009, *Natur*, **462**, 624
- Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, *ApJ*, **689**, 1044
- Gies, D. R., & Bolton, C. T. 1986, *ApJS*, **61**, 419
- Glebbeek, E., Gaburov, E., Portegies Zwart, S., & Pols, O. R. 2013, *MNRAS*, **434**, 3497
- Gvaramadze, V. V., & Bomans, D. J. 2008a, *A&A*, **485**, L29
- Gvaramadze, V. V., & Bomans, D. J. 2008b, *A&A*, **490**, 1071
- Gvaramadze, V. V., Kniazev, A. Y., Kroupa, P., & Oh, S. 2011, *A&A*, **535**, A29
- Habibi, M., Stolte, A., Brandner, W., Hußmann, B., & Motohara, K. 2013, *A&A*, **556**, A26
- Harfst, S., Portegies Zwart, S., & Stolte, A. 2010, *MNRAS*, **409**, 628
- Hénault-Brunet, V., Evans, C. J., Sana, H., et al. 2012, *A&A*, **546**, A73
- Herrero, A., Corral, L. J., Villamariz, M. R., & Martín, E. L. 1999, *A&A*, **348**, 542
- Hillenbrand, L. A., Massey, P., Strom, S. E., & Merrill, K. M. 1993, *AJ*, **106**, 1906
- Hills, J. G., & Day, C. A. 1976, *ApL*, **17**, 87
- Hußmann, B., Stolte, A., Brandner, W., Gennaro, M., & Liermann, A. 2012, *A&A*, **540**, A57
- Hunter, D. A., Shaya, E. J., Holtzman, J. A., et al. 1995, *ApJ*, **448**, 179
- Hunter, D. A., Vacca, W. D., Massey, P., Lynds, R., & O’Neil, E. J. 1997, *AJ*, **113**, 1691
- Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, *MNRAS*, **315**, 543
- Hurley, J. R., Tout, C. A., Aarseth, S. J., & Pols, O. R. 2001, *MNRAS*, **323**, 630
- Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, *MNRAS*, **329**, 897
- Izzard, R. G., Dray, L. M., Karakas, A. I., Lugaro, M., & Tout, C. A. 2006, *A&A*, **460**, 565
- Izzard, R. G., Glebbeek, E., Stancliffe, R. J., & Pols, O. R. 2009, *A&A*, **508**, 1359
- Izzard, R. G., Tout, C. A., Karakas, A. I., & Pols, O. R. 2004, *MNRAS*, **350**, 407
- Kim, S. S., Figer, D. F., Kudritzki, R. P., & Najarro, F. 2006, *ApJL*, **653**, L113
- Kippenhahn, R., & Meyer-Hofmeister, E. 1977, *A&A*, **54**, 539
- Koen, C. 2006, *MNRAS*, **365**, 590
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Kudryavtseva, N., Brandner, W., Gennaro, M., et al. 2012, *ApJL*, **750**, L44
- Langer, N. 2009, *Natur*, **462**, 579
- Langer, N. 2012, *ARA&A*, **50**, 107
- Langer, N., Norman, C. A., de Koter, A., et al. 2007, *A&A*, **475**, L19
- Larson, R. B., & Starrfield, S. 1971, *A&A*, **13**, 190
- Lejeune, T., & Schaerer, D. 2001, *A&A*, **366**, 538
- Liermann, A., Hamann, W.-R., & Oskinova, L. M. 2009, *A&A*, **494**, 1137
- Liermann, A., Hamann, W.-R., & Oskinova, L. M. 2012, *A&A*, **540**, A14
- Liermann, A., Hamann, W.-R., Oskinova, L. M., Todt, H., & Butler, K. 2010, *A&A*, **524**, A82
- Lim, B., Chun, M.-Y., Sung, H., et al. 2013, *AJ*, **145**, 46
- Maíz Apellániz, J., & Úbeda, L. 2005, *ApJ*, **629**, 873
- Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, **482**, 883
- Martins, F., Hillier, D. J., Paumard, T., et al. 2008, *A&A*, **478**, 219
- Massey, P. 2003, *ARA&A*, **41**, 15
- Massey, P., & Hunter, D. A. 1998, *ApJ*, **493**, 180
- McCrea, W. H. 1964, *MNRAS*, **128**, 147
- Negueruela, I., Marco, A., Herrero, A., & Clark, J. S. 2008, *A&A*, **487**, 575
- Neo, S., Miyaji, S., Nomoto, K., & Sugimoto, D. 1977, *PASJ*, **29**, 249
- Nieuwenhuijzen, H., & de Jager, C. 1990, *A&A*, **231**, 134
- Nishiyama, S., Tamura, M., Hatano, H., et al. 2009, *ApJ*, **696**, 1407
- Oey, M. S., & Clarke, C. J. 2005, *ApJL*, **620**, L43
- Öpik, E. 1924, *PTarO*, **25**, 1
- Pang, X., Grebel, E. K., Allison, R. J., et al. 2013, *ApJ*, **764**, 73
- Pang, X., Grebel, E. K., & Altmann, M. 2010, in IAU Symp. 266, *Star Clusters: Basic Galactic Building Blocks throughout Time and Space*, ed. R. de Grijs & J. R. D. Lépine (Cambridge: Cambridge Univ. Press), 24
- Perets, H. B., & Šubr, L. 2012, *ApJ*, **751**, 133
- Pols, O. R., & Marinus, M. 1994, *A&A*, **288**, 475
- Pols, O. R., Schroder, K.-P., Hurley, J. R., Tout, C. A., & Eggleton, P. P. 1998, *MNRAS*, **298**, 525
- Portegies Zwart, S., Gaburov, E., Chen, H.-C., & Gürkan, M. A. 2007, *MNRAS*, **378**, L29
- Portegies Zwart, S. F., Makino, J., McMillan, S. L. W., & Hut, P. 1999, *A&A*, **348**, 117
- Ramírez-Agudelo, O. H., Simón-Díaz, S., Sana, H., et al. 2013, *A&A*, **560**, A29
- Rieke, G. H., & Lebofsky, M. J. 1985, *ApJ*, **288**, 618
- Sabbi, E., Lennon, D. J., Gieles, M., et al. 2012, *ApJL*, **754**, L37
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Sana, H., de Koter, A., de Mink, S. E., et al. 2013, *A&A*, **550**, A107
- Sana, H., de Mink, S. E., de Koter, A., et al. 2012, *Sci*, **337**, 444
- Sandage, A. R. 1953, *AJ*, **58**, 61
- Schneider, F. R. N., Izzard, R. G., Langer, N., & de Mink, S. E. 2013, *ApJ*, submitted
- Schnurr, O., Casoli, J., Chené, A.-N., Moffat, A. F. J., & St-Louis, N. 2008, *MNRAS*, **389**, L38
- Stolte, A., Brandner, W., Grebel, E. K., Lenzen, R., & Lagrange, A.-M. 2005, *ApJL*, **628**, L113
- Stolte, A., Grebel, E. K., Brandner, W., & Figer, D. F. 2002, *A&A*, **394**, 459
- Stone, R. C. 1991, *AJ*, **102**, 333
- Ulrich, R. K., & Burger, H. L. 1976, *ApJ*, **206**, 509
- van Bever, J., & Vanbeveren, D. 1998, *A&A*, **334**, 21
- Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2000, *A&A*, **362**, 295
- Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2001, *A&A*, **369**, 574
- Weidner, C., & Kroupa, P. 2004, *MNRAS*, **348**, 187
- Wellstein, S., Langer, N., & Braun, H. 2001, *A&A*, **369**, 939